

MENINGKATKAN KINERJA INDIVIDU MELALUI KRITIK/SARAN MENGUNAKAN RECOMMENDER SYSTEM

Zen Munawar

Universiti Teknikal Malaysia Melaka

E-mail:munawarzen@gmail.com

Abstrak : *Recommender system*. sebagai bagian yang terintegrasi akan memberikan penyesuaian personalisasi individu dalam pengalaman menggunakan web. Penggalan pengetahuan produk dapat memandu pengguna saat melakukan pencarian suatu produk. Memberikan bantuan dan kinerja pengguna yang cerdas tugas pada nama pengguna akan membutuhkan pemahaman tentang tujuan dan preferensi mereka. Dengan demikian, umpan balik pengguna mempunyai peran penting dalam arti bahwa hal itu membantu mengarahkan pencarian ke arah rekomendasi yang "baik". Idealnya, sistem harus mampu menafsirkan umpan balik yang diberikan pengguna secara efektif, dan kemudian merespons dengan memberikan serangkaian rekomendasi yang "lebih baik". Dalam penelitian ini bertujuan menyelidiki bentuk umpan balik yang dikenal sebagai kritik/saran. Meskipun sejumlah besar pemberi rekomendasi sangat cocok dengan bentuk umpan balik ini, penulis berpendapat bahwa dengan sendirinya hal itu dapat menyebabkan dialog rekomendasi yang tidak efisien. Sebagai solusi, diusulkan teknik rekomendasi baru yang memiliki kemampuan yang secara dramatis meningkatkan kegunaan dari kritik/saran.

Kata Kunci : Preferensi, kinerja, kritik, *Recommender system*

1. Pendahuluan

Sistem pemberi rekomendasi yang dipersonalisasi menggabungkan ide dari pencarian informasi, Kecerdasan buatan (KB), pemodelan pengguna dan desain antarmuka untuk menyediakan pengguna dengan lingkungan yang lebih adaptif untuk pengambilan informasi. Sistem rekomendasi reaktif dirancang untuk membuat rekomendasi berdasarkan permintaan pengguna; misalnya, Entree (R. Burke, 1997) membuat rekomendasi restoran berdasarkan permintaan yang menentukan fitur seperti jenis masakan, biaya, dll. Sebaliknya, pemberi rekomendasi proaktif beroperasi tanpa perlunya permintaan eksplisit, secara proaktif mendorong rekomendasi kepada pengguna; misalnya, PTVPlus membuat saran program TV kepada pengguna berdasarkan preferensi menonton yang mereka pelajari (B. Smyth and P. Cotter, 2000). Umpan balik pengguna adalah komponen vital dari sebagian besar pemberi rekomendasi, yang memungkinkan mereka untuk beradaptasi secara tepat dengan kebutuhan pengguna target, dan menetapkan sistem pemberi rekomendasi selain dari sistem pencarian informasi yang lebih tradisional. Sampai saat ini para peneliti telah fokus pada elisitasi nilai dan umpan balik berbasis peringkat, tetapi baru-baru ini dua alternatif, kritik (atau tweaker) dan umpan balik berbasis preferensi, telah menjadi semakin relevan dalam banyak skenario aplikasi penting. Dalam penelitian ini fokus pada

rekomendator reaktif yang menggunakan kritik sebagai sumber umpan balik utama mereka, sebagaimana dicontohkan oleh apa yang disebut sistem rekomendator FindMe (R. Burke, 1997) (R. Burke, 2000). Misalnya, di Ent Restaurant Recommender, pengguna cenderung memiliki pengetahuan terbatas tentang restoran yang direkomendasikan, tetapi memang demikian mampu mengenali 'apa yang mereka inginkan ketika mereka melihatnya'.



Gambar 1 Menunjukkan Layar Awal Pada Awal Sesi Baru.

Pengguna memiliki dua opsi: (1) mereka dapat menentukan preferensi restoran mereka dengan memilih fitur tertentu (misalnya, lokasi pilihan, masakan, suasana, harga, dll.) dari daftar drop-down yang disediakan, atau (2) tentukan nama restoran yang sudah mereka kenal dan sukai, dan gunakan deskripsinya sebagai queri. Ketika pengguna mengirimkan queri mereka ke mesin pemberi rekomendasi Ent, sistem merespons dengan menghadirkan kepada pengguna deskripsi restoran paling mirip yang tersedia (lihat Gambar 2).



Gambar 2. Ketika sebuah rekomendasi dikembalikan pengguna dapat menerapkan salah satu dari serangkaian kritik yang telah ditentukan sebelumnya.

Pengguna kemudian dapat menerima rekomendasi ini, atau mengkritiknya dengan meminta yang lain yang berbeda terkait dengan salah satu dari tujuh fitur yang telah ditentukan sebelumnya (misal., Lebih murah, lebih bagus, lebih tenang, dll.). Gambar 3 menunjukkan bagaimana pengguna akan memilih kritik atas fitur masakan. Kritik yang diterapkan kemudian bertindak sebagai filter dalam proses pengambilan, sehingga setiap kandidat restoran yang tidak memenuhi kritik tidak dipertimbangkan untuk siklus rekomendasi berikutnya. Yang penting, setelah halaman pertama, tidak ada tindakan yang membutuhkan lebih dari satu klik, dan pengguna mengendalikan arah pencarian.



Gambar 3. Contoh kritik. Pengguna ingin melihat rekomendasi lain seperti yang sekarang, tetapi dengan masakan yang berbeda.

Khususnya, mengkritik keseimbangan menarik antara upaya bahwa pengguna harus berinvestasi dalam memberikan umpan balik dan konten informasi umpan balik; lihat Bagian 2. Kritik relatif mudah diterapkan dan mereka tidak mengharuskan pengguna untuk memiliki pengetahuan domain terperinci, namun mereka memberikan informasi yang berharga bagi pemberi rekomendasi. berpendapat bahwa ini menjadikan kritik sebagai bentuk umpan balik yang sangat berguna untuk banyak aplikasi praktis. Namun, juga menyoroti bagaimana kritik itu sendiri dapat mengarah pada dialog rekomendasi yang berlarut-larut. Sebagai solusi, menggambarkan dan mengevaluasi pendekatan baru untuk rekomendasi, yang disebut seleksi adaptif, yang dapat digunakan secara signifikan meningkatkan kinerja. jumlah rekomendasi yang menggunakan kritik, memperpendek dialog rekomendasi hingga 60%. Sebelum melanjutkan, perlu dicatat bahwa dalam penelitian terkait telah menunjukkan bagaimana pendekatan ini juga dapat digunakan dengan bentuk lain dari umpan balik seperti umpan balik berbasis preferensi (L. McGinty and B. Smyth, 2003) (L. McGinty and B. Smyth, 2002) (L. McGinty and B. Smyth, 2003).

2. Background

Sistem rekomendasi berbasis konten yang reaktif biasanya mengadopsi gaya interaksi percakapan dengan pengguna, memilih item untuk rekomendasi dan memperoleh umpan balik dari pengguna sebelum membuat kumpulan rekomendasi berikutnya (D.W. Aha, 2000) (R. Burke, 1997) (R. Burke, 2000) M. Goker and C. Thompson. (2000) (S. Rogers, 1997) (Shimazu. ExpertClerk, 2001). Umpan balik memungkinkan rekomendasi untuk membuat saran yang lebih baik dengan

mengadaptasi pemahamannya saat ini tentang persyaratan pengguna. terutama tertarik dalam mengembangkan produk rekomendasi untuk pasar konsumen, dan untuk digunakan di berbagai perangkat, dari PC tradisional ke perangkat yang jauh lebih terbatas seperti ponsel yang mendukung Internet. Metode umpan balik yang ideal harus berbiaya rendah dan memberikan informasi yang tidak ambigu mengenai niat pengguna. Itu juga harus berlaku di berbagai jenis pengguna, bahkan yang dengan keahlian domain terbatas. Akhirnya, mungkin untuk menangkap umpan balik pengguna pada perangkat yang paling terbatas sekalipun.

2.1 Umpan Balik Pengguna dalam Sistem Rekomendasi Personalisasi

Seperti disebutkan di atas, kita dapat dengan mudah mengkategorikan strategi umpan balik dalam hal: biaya bagi pengguna yang memberikan umpan balik; tingkat ambiguitas yang melekat dalam umpan balik; tingkat keahlian domain yang dibutuhkan oleh pengguna untuk memberikan umpan balik; dan jenis antarmuka pengguna yang diperlukan untuk menangkap umpan balik; lihat Gambar 4. Sebagai contoh, nilai elisitasi (mis. "Saya ingin PC 2GHz"), mungkin bentuk umpan balik yang paling umum, adalah sumber informasi yang kaya; mengetahui bahwa pengguna tertarik pada item dengan fitur tertentu memungkinkan merekomendasikan untuk menghilangkan banyak item yang tidak relevan dari pertimbangan.

Umpan Balik	Biaya	Ambigu	Keahlian	Antarmuka
Elisitasi Nilai	***	*	***	***
Peringkat	**	***	**	*
Kritik	**	**	**	*
Pilihan	*	***	*	*

Kunci : * Rendah ** Moderat *** Tinggi

Gambar 4. Perbandingan strategi umpan balik

Namun, untuk memberikan umpan balik ini, pengguna memerlukan keahlian domain tingkat tinggi untuk dapat menentukan nilai fitur yang masuk akal. Selain itu, pengguna harus bersedia menjawab pertanyaan langsung dan khusus yang diajukan oleh para pemberi rekomendasi. Terakhir, memberikan umpan balik tingkat fitur yang terperinci menuntut antarmuka pengguna yang canggih (mis. Entri teks, kotak centang, daftar drop), sehingga membatasi penggunaannya pada beberapa perangkat seluler, misalnya. Sebaliknya, umpan balik berbasis preferensi (mis. "Saya lebih suka PC1") adalah bentuk umpan balik berbiaya rendah dan dapat disediakan oleh pengguna, melalui antarmuka yang sederhana, dan hanya dengan pemahaman dasar domain. Itu berbiaya rendah karena pengguna dapat menunjukkan preferensi mereka dalam satu klik. Bandingkan dengan elisitasi nilai di mana pengguna harus memilih fitur spesifik dari kasus tertentu dan memberikan nilai spesifik untuk fitur itu. Namun, dengan sendirinya preferensi sederhana untuk suatu item secara inheren ambigu sehubungan dengan maksud pengguna. Umpan balik berbasis preferensi tidak memberikan informasi tentang

fitur tertentu yang memotivasi preferensi atau pilihan pengguna. Dengan demikian hanya memiliki kapasitas terbatas untuk memandu proses rekomendasi; lihat (L. McGinty and B. Smyth, 2002) (L. McGinty and B. Smyth, 2003) untuk kemajuan terbaru dengan bentuk umpan balik ini.

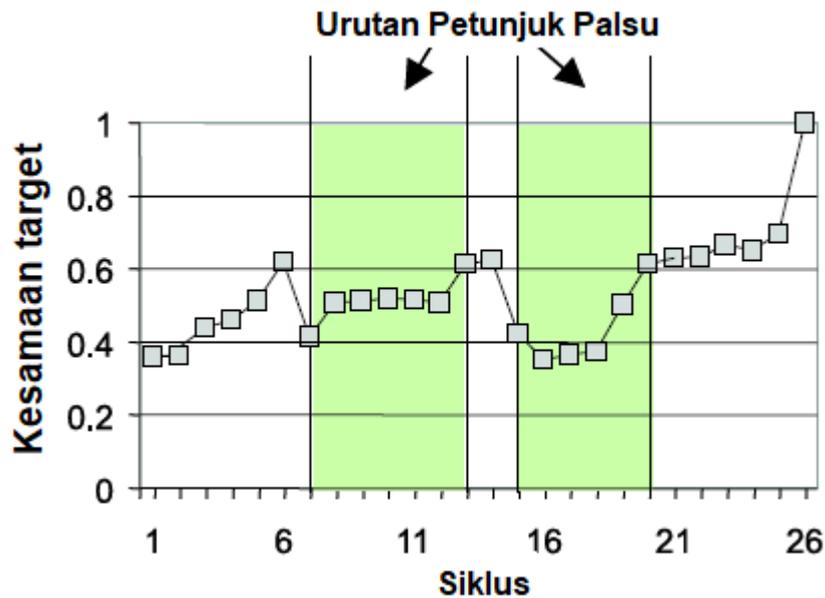
2.2 Jebakan Kritik

Baik elisitasi nilai maupun umpan balik berbasis preferensi adalah bentuk umpan balik yang ideal (yaitu berbiaya rendah, tidak ambigu, keahlian minimal, dan antarmuka yang sederhana), tetapi kritik tampaknya sudah dekat. Ini adalah kasus khusus dari elisitasi nilai - alih-alih memberikan nilai fitur tertentu, pengguna mengindikasikan kritik fitur (atau tweak) untuk membatasi rentang nilai fitur tersebut. Sebagai contoh, tweak, "Tunjukkan saya lebih seperti PC1 tapi lebih murah", memungkinkan pemberi rekomendasi untuk fokus pada produk yang lebih murah daripada PC1 dalam siklus rekomendasi berikutnya tanpa mengharuskan pengguna untuk memberikan rincian harga spesifik.

Manfaat mengkritisi sangat menarik. Ini memberikan indikasi yang relatif jelas tentang kebutuhan pengguna saat ini. Ini berbiaya rendah, mudah diimplementasikan dengan antarmuka yang paling sederhana, dan bahkan dapat diterapkan oleh pengguna dengan hanya pemahaman yang moderat tentang ruang produk. Namun, masih ada beberapa kekhawatiran. Ada relatif sedikit pekerjaan evaluasi yang dilakukan pada kritik, terutama sehubungan dengan dampaknya pada efisiensi rekomendasi. Seberapa efektif kritik individu dalam memfokuskan dialog rekomendasi? Dipersonalisasi sistem rekomendasi yang memanfaatkan umpan balik berbasis preferensi memiliki kecenderungan untuk menyesuaikan dengan preferensi saat ini akan menyesatkan rekomendasi untuk mengikuti petunjuk palsu ke dalam bagian ruang item yang tidak relevan [8]. Apakah masalah ini ada dengan mengkritisi dan jika demikian bagaimana cara mengatasinya? Pertimbangkan sistem pemberi rekomendasi yang menyajikan kepada pengguna serangkaian 3 rekomendasi selama setiap siklus rekomendasi dan mengharapkan pengguna untuk memandu siklus rekomendasi berikutnya dengan mengkritisi salah satu fitur dari item yang disukai. Gambar 5 menggambarkan profil kemiripan rekomendasi untuk permintaan tunggal yang diambil dari domain Perjalanan. Grafik memplot kesamaan antara item yang dikritik dan beberapa target ideal yang telah ditetapkan. Di sini terlihat bahwa terdapat 26 siklus diperlukan untuk menemukan item target untuk permintaan ini. Demi kejelasan maka telah dipilih permintaan yang sangat sulit yang menghasilkan dialog rekomendasi yang berlarut-larut. Jelas tidak mungkin bahwa pengguna akan mentolerir dialog yang panjang; pada kenyataannya sangat mungkin bahwa sebagian besar pengguna akan menyerah setelah 5 hingga 10 siklus. Namun demikian dialog panjang seperti itu dapat terjadi dengan mengkritik sebagai sumber umpan balik.

Melihat profil kesamaan ini secara lebih rinci menunjukkan bahwa kemajuan yang baik dibuat menuju target selama 6 siklus pertama, seperti yang ditunjukkan oleh meningkatnya kesamaan kasus yang disukai dengan target. Namun, rekomendasi

yang dibuat dalam siklus ke-7 adalah petunjuk palsu karena tidak ada satupun yang lebih mendekati masalah target daripada kasus yang disukai dari siklus sebelumnya. Bahkan yang terbaik dari kesalahan ini, yang merupakan kasus preferensi untuk siklus ke-7, memiliki kesamaan target saja sebesar 0,4 yang kurang dari kesamaan target 0,6 untuk kasus preferensi sebelumnya; pada dasarnya serangkaian rekomendasi ini telah menyebabkan pengguna tersesat.



Gambar 5. Contoh profil kemiripan sesi rekomendasi untuk sampel query .

Petunjuk palsu lainnya diikuti dari siklus 15 dan, secara keseluruhan, petunjuk palsu ini berkontribusi 11 siklus tambahan untuk sesi yang menghasilkan dialog rekomendasi panjang yang tidak tidak menjamin untuk memberikan rekomendasi yang semakin relevan kepada pengguna. telah menemukan jenis profil kesamaan ini menjadi umum di rekomendasi berdasarkan kritik, terutama untuk pertanyaan yang awalnya samar atau sulit untuk dipenuhi. Dan, kecuali petunjuk palsu dapat dideteksi dan dikendalikan, masalah ini cenderung membatasi keberhasilan praktis sistem rekomendasi tersebut.

3. Mengkritik dalam Rekomendasi Berbasis Perbandingan

Rekomendasi berbasis perbandingan (L. McGinty and B. Smyth, 2002) adalah kerangka kerja untuk sistem rekomendasi berbasis konten yang reaktif yang menekankan pentingnya umpan balik selama setiap siklus rekomendasi. Rekomendasi berbasis perbandingan pada awalnya diusulkan sebagai test-bed untuk strategi umpan balik berbasis preferensi tetapi sama berlaku untuk bentuk umpan balik lain, terutama kritik. Karenanya, dalam tulisan ini menyelidiki penggunaannya dengan mengkritisi, dan mengusulkan dan mengevaluasi selection (AS) sebagai strategi rekomendasi yang mampu menangani kesalahan-lead masalah...

3.1 Review

Rekomendasi berbasis perbandingan adalah proses rekomendasi percakapan dimana pengguna terlibat dalam dialog yang diperluas yang terdiri dari urutan pemulihan siklus rekomendasi. Algoritma dasar (Penjelasan terdapat pada Gambar 6) terdiri dari 3 langkah utama: (1) item baru direkomendasikan kepada pengguna berdasarkan permintaan saat ini; (2) pengguna meninjau rekomendasi dan menunjukkan kasus preferensi sebagai umpan balik; (3) informasi tentang perbedaan antara item yang dipilih dan alternatif yang tersisa digunakan untuk merevisi permintaan untuk siklus berikutnya (L. McGinty and B. Smyth, 2002).

```

1.  define Comparison-Based-Recommend(q, CB, k)
2.     $i_{p-1}, i_p \leftarrow \text{null}$ 
3.    do
4.       $R \leftarrow \text{ItemRecommend}(q, CB, c, k, i_p, i_{p-1})$ 
5.       $\langle i_p, c \rangle \leftarrow \text{UserReview}(R, CB)$ 
6.       $Q \leftarrow \text{QueryRevise}(q, i_p, R)$ 
7.       $i_{p-1} \leftarrow i_p$ 
8.    until UserAccepts( $i_p$ )

9.  define QueryRevise(q,  $i_p$ , R)
10.    $R' \leftarrow R - \{i_p\}$ 
11.    $q \leftarrow i_p$ 
12.   return q

13. define UserReview(R, CB)
14.    $i_p \leftarrow \text{user's preferred case from } R$ 
15.    $c \leftarrow \text{user critique for some } f \in i_p$ 
16.    $CB \leftarrow CB - R$ 
17.   return  $\langle i_p, c \rangle$ 

18. define ItemRecommend(q, CB, c, k,  $i_p, i_{p-1}$ )
19.    $CB' \leftarrow \{i \in CB \mid \text{Satisfies}(i, c)\}$ 
20.   if ( $i_p \neq \text{null}$ ) && ( $i_p == i_{p-1}$ )
21.      $CB'' \leftarrow \text{sort } CB' \text{ in by decreasing sim to } q$ 
22.      $R \leftarrow \text{top } k \text{ items in } CB''$ 
23.   else  $R \leftarrow \text{BoundedGreedySelection}(q, CB, k, b)$ 
24.   return R

25. define BoundedGreedySelection (q, CB, k, b)
26.    $CB' := \text{bk items in } CB \text{ that are most similar to } q$ 
27.    $R := \{\}$ 
28.   For j := 1 to k
29.     Sort  $CB'$  by  $\text{Quality}(q, i, R)$  for each case  $i$  in  $CB'$ 
30.      $R := R + \text{First}(CB')$ 
31.      $CB' := CB' - \text{First}(CB')$ 
32.   EndFor
33.   return R

... where  $\text{Quality}(q, i, R) = \alpha \cdot \text{Sim}(q, I) + (1 - \alpha) \cdot \text{RelDiv}(i, R)$ 
        $\text{RelDiv}(i, R) = \sum_{j=1..m} (1 - \text{Sim}(i, r_j)) / m$  if  $R = \{\}$ ; = 1 otherwise

```

Gambar 6. Algoritma rekomendasi berbasis perbandingan disesuaikan untuk kritik dengan seleksi adaptif.

Proses rekomendasi berakhir baik ketika pengguna disajikan dengan item yang cocok atau ketika mereka menyerah. Biasanya langkah ItemRecommend (baris 4 pada Gambar 6) mengambil k item yang paling mirip dengan query, pengguna menunjukkan case dari antara item yang direkomendasikan yang paling sesuai dengan kebutuhan mereka pada langkah UserReview (baris 5 pada Gambar 6), dan Langkah QueryRevise (baris 6 dari Gambar 6) berfokus pada pembaruan query saat ini berdasarkan apa yang dapat dipelajari dari pengguna umpan balik. Pekerjaan sebelumnya telah memeriksa sejumlah strategi pembaruan query untuk digunakan dengan umpan balik berbasis preferensi (mis. "tunjukkan lebih banyak seperti ini (MLT)") di mana query diuraikan dengan fitur dari case yang disukai yang terbaik mencerminkan preferensi implisit pengguna (L. McGinty and B. Smyth, 2002).

3.2 Menyesuaikan Rekomendasi Berbasis Perbandingan untuk Kritik

Menyesuaikan rekomendasi berdasarkan perbandingan untuk mengkritisi relatif lebih mudah. Misalnya, dalam domain Perjalanan anggap pengguna direkomendasikan liburan dua minggu seharga \$ 2000 di Venesia, Italia di hotel bintang tiga, tetapi menunjukkan bahwa mereka mencari sesuatu yang serupa tetapi lebih murah (kritik $< \$ 2000$ pada fitur harga). Selama siklus berikutnya hanya barang-barang yang harganya lebih murah dari \$ 2000 akan dipertimbangkan untuk seleksi. Dengan demikian ItemRecommend menjadi proses dua langkah. Pada langkah pertama, semua item yang gagal memenuhi kritik (itu dengan fitur harga $\geq \$ 2000$) dihilangkan. Pada langkah kedua, item yang tersisa diberi peringkat berdasarkan kesamaannya dengan query yang diperbarui, dan k teratas adalah terpilih. Perhatikan bahwa dalam hal ini akan menganggap bahwa revisi query mengikuti lebih sederhana seperti ini (MLT) strategi di mana item yang disukai saat ini digunakan sebagai query baru dalam siklus rekomendasi berikutnya (L. McGinty and B. Smyth, 2002). memanfaatkan empat tipe dasar kritik: $<, >, =, \text{ dan } =$. Jelas, $<\text{dan}>$ hanya bisa diterapkan pada fitur numerik.

3.3 Seleksi Adaptif

Seperti yang disebutkan sebelumnya, strategi pemilihan item standar memilih k item yang paling mirip dengan permintaan, tetapi karya terbaru menunjukkan bahwa faktor selain kesamaan harus dipertimbangkan selama pemilihan. Penelitian telah menunjukkan bahwa berbasis kesamaan metode cenderung mengurangi keragaman rekomendasi jika item yang dipilih mirip satu sama lain, serta mirip dengan permintaan (D. Bridge, 2001) (D. McSherry, 2002) (Shimazu. ExpertClerk, 2001) (H. Shimazu, A. Shibata, and K. Nihei. ExpertGuide, 2002) B. (Smyth and P. McClave, 2001). Dengan meningkatkan tingkat keragaman di antara item yang direkomendasikan, dimungkinkan untuk mencakup lebih banyak

ruang item, dalam siklus tertentu, dan sebagai konsekuensinya mungkin meningkatkan efisiensi rekomendasi.

Sejumlah mekanisme peningkatan keragaman telah diusulkan dalam literatur, tetapi teknik ini memperkenalkan keanekaragaman dalam atau cara yang seragam. Sementara peningkatan keragaman seleksi dapat menawarkan an merekomendasikan peningkatan kemampuan untuk menutupi ruang item, hal itu dapat menyebabkan inefisiensi baru (L. McGinty and B. Smyth, 2003). Oleh desain, metode seleksi peningkatan keragaman melewati item tertentu yang mungkin jika tidak dipilih, dan ini menimbulkan masalah jika target kebetulan salah satu dari barang yang dilewati ini; dalam hal ini item target yang benar telah dilewati lebih dalam pencarian untuk meningkatkan keragaman rekomendasi. Yang lebih fleksibel mekanisme untuk memperkenalkan keragaman tampaknya tepat. Jika yang merekomendasikan adalah berfokus pada wilayah masalah target, kesamaan harus ditekankan, kalau tidak keragaman harus memainkan peran. Namun demikian, hal itu tidak akan memperjelas bagaimana menilai ketika pemberi rekomendasi difokuskan di wilayah target. Gagasan kunci dalam seleksi adaptif (AS) adalah dengan menentukan apakah kebanyakan rekomendasi baru-baru ini menunjukkan peningkatan pada yang dibuat dalam siklus sebelumnya, mungkin untuk menilai apakah rekomendasi tersebut benar-benar fokus dalam pencariannya, dan dengan demikian apakah keragaman harus dimasukkan ke dalam pemilihan item. Ini dicapai dengan membuat dua modifikasi pada strategi pemilihan item berbasis kesamaan standar (misalnya. Metode Irecomendasi item yang diuraikan dalam baris 22-27 pada Gambar 6).

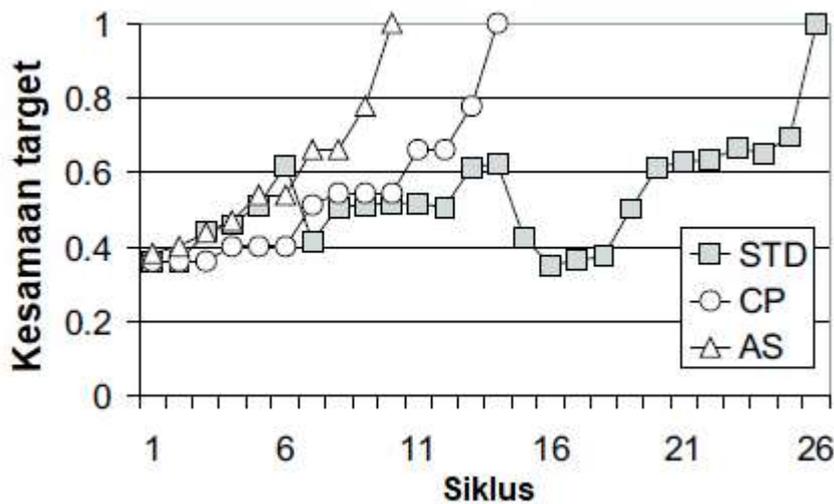
Pertama, alih-alih membuat k rekomendasi baru di setiap siklus baru, item preferensi saat ini (item yang dikritik) ditambahkan ke $k - 1$ rekomendasi baru; menyebutnya sebagai membawa preferensi (CP). Modifikasi ini sendiri menunjukkan redundansi, dalam arti bahwa item yang sebelumnya dilihat diulang dalam satu atau lebih siklus di masa depan. Namun, termasuk preferensi sebelumnya memungkinkan untuk menghindari masalah kesalahan yang timbul dan telah disebutkan di atas.

Jika tidak ada $k - 1$ item baru yang relevan maka pengguna dapat terus mengkritik preferensi yang dibawa preferensi bukannya dipaksa untuk mengkritik rekomendasi yang kurang relevan. Sebagai contoh, profil kesamaan untuk metode ini termasuk dalam Gambar 7, dengan 14 siklus diperlukan, dan sejumlah kesamaan dataran tinggi dalam bukti (siklus 4-6, 8-10 dan 11-12) ketika pengguna memilih kembali item preferensi. Jika pengguna mengkritik preferensi yang dibawa maka itu menunjukkan bahwa item k items 1 lainnya kurang relevan daripada item yang diangkat, dan dengan demikian pemberi rekomendasi gagal membuat kemajuan positif menuju target. Jika pengguna memilih untuk mengkritik salah satu item yang baru direkomendasikan, maka dari itu harus karena lebih dekat ke target, dan dengan demikian kemajuan positif telah dibuat. Bagian kedua dari strategi seleksi adaptif akan mengambil keuntungan dari ide ini dengan memperkenalkan keragaman ke dalam putaran seleksi berikutnya jika dan hanya jika pengguna telah mengkritik preferensi yang dibawa. Profil kesamaan

untuk seleksi adaptif juga disajikan pada Gambar 7. Sekali lagi penurunan yang signifikan dalam jumlah siklus sangat jelas, tanpa ada penurunan kesamaan target; saat ini hanya diperlukan 10 siklus, dengan preferensi yang dibawa terpilih kembali dalam siklus 6 dan 8.

4. Evaluasi Eksperimental

Keberhasilan setiap penasehat percakapan sangat tergantung pada efisiensi dialognya. Secara umum, dialog pendek yang menghadirkan rekomendasi yang semakin baik kepada pengguna lebih cenderung menghasilkan kesuksesan yang lebih besar daripada sebelumnya (M. Doyle and P. Cunningham, 2000) (D. McSherry, 2001). Pada bagian ini menjelaskan hasil dari evaluasi komprehensif pengkritikan yang berfokus pada manfaat mengkritisi dengan membawa preferensi dan seleksi adaptif baru teknik lebih dari kritikan standar.



Gambar. 7. Preferensi profil kesamaan kasus untuk tiga versi yang berbeda dari strategi kritik; standar mengkritisi (STD), mengkritisi dengan membawa preferensi (CP), dan mengkritisi dengan seleksi adaptif (AS). yang lebih panjang

4.1 Asumsi Dasar

berasumsi bahwa pengguna memiliki beberapa persyaratan produk dan mereka mencari produk yang memenuhi persyaratan ini. Namun, lebih jauh menganggap bahwa persyaratan ini cukup kabur dan bahwa pengguna tidak bisa hanya daftar fitur dan cari dari sana. Sebaliknya, pengguna menyediakan beberapa fitur awal dan kemudian dibantu untuk menemukan produk yang baik dengan mengkritik fitur tambahan dari rekomendasi perantara. Rekomendasi perantara ini membantu untuk mempengaruhi persyaratan pengguna dan menguraikan permintaan awal mereka.

4.2 Setup

Dalam evaluasi ini membandingkan kinerja tiga teknik mengkritisi dalam merekomendasikan berbasis perbandingan (dengan $k = 3$):

1. STD menggunakan kritik dasar dan merupakan tolok ukur ,
2. CP mengimplementasikan preferensi pembawa, dan
3. AS mengimplementasikan seleksi adaptif.

Kumpulan Data Dua kumpulan data publik, dari domain Travel dan PC digunakan untuk evaluasi. Basis kasus Perjalanan yang akrab berisi 1024 kasus, masing-masing menggambarkan liburan tertentu dalam hal fitur seperti wilayah, durasi, akomodasi, harga dll (lihat Gambar 8). Basis kasus PC berisi 120 kasus, menggambarkan PC yang unik dalam hal fitur seperti pabrikan, tipe prosesor, kecepatan prosesor, dll (lihat Gambar 9).



Jenis Liburan	Bathing
Harga (\$)	2498
# orang	2
Negara	Mesir
Transportasi	Penerbangan
Durasi (hari)	10
Sesi (bulan)	April
Akomodasi	Bintang dua
Hotel	Hotel White House

Gambar 8. Contoh Kasus Travel

Metodologi Menggunakan metodologi satu-keluar, setiap item (basis) dari set data dihapus sementara dan digunakan dalam dua cara. Pertama, berfungsi sebagai dasar untuk seperangkat pertanyaan yang dibangun dengan mengambil himpunan bagian acak fitur item. Kedua, memilih item yang paling mirip dengan basis aslinya. Item ini berfungsi sebagai target rekomendasi untuk percobaan. Dengan demikian, basis mewakili permintaan ideal untuk pengguna, permintaan yang dihasilkan adalah permintaan awal yang diberikan pengguna kepada pemberi rekomendasi, dan target adalah item terbaik yang tersedia untuk pengguna berdasarkan ideal mereka. Setiap query yang dihasilkan adalah masalah pengujian untuk pemberi rekomendasi, dan dalam setiap siklus rekomendasi tweak diterapkan ke item yang paling mirip dengan item target yang diketahui; tweak didasarkan pada fitur acak dari item ini. Untuk setiap kumpulan data, tiga kelompok query yang berbeda dihasilkan dari berbagai tingkat kesulitan (mudah,

sedang, sulit); kesulitan didasarkan pada jumlah siklus yang dibutuhkan oleh STD.



Pabrik	Compaq
Tipe Prosesor	Intel I7
Kecepatan Prosesor (Ghz)	3,4
Monitor (Inchi)	14
Tipe	Laptop
Memory (Gb)	16
Hardisk (Tb)	1
Harga (\$)	2300

Gambar 9. Contuh Kasus PC

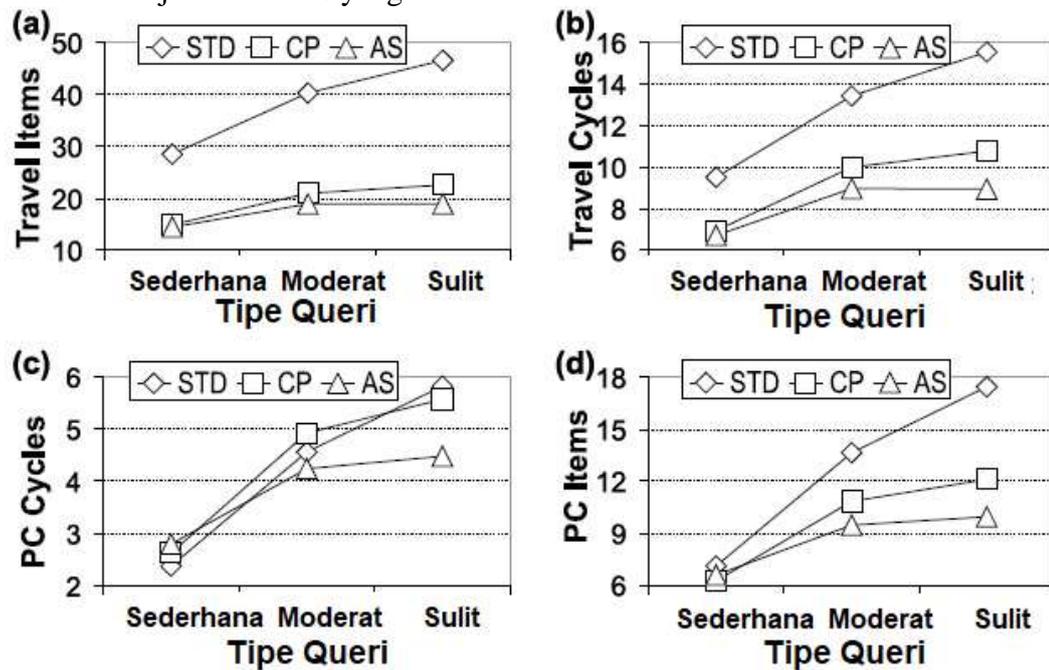
Karena perbedaan kompleksitas antara PC dan data Perjalanan menetapkan ambang siklus yang sesuai dengan tingkat kesulitan tertentu tidak sama untuk kedua kumpulan data dan, secara umum, queri PC lebih sulit daripada kueri Perjalanan.

4.3 Efisiensi Rekomendasi

Untuk menguji efisiensi rekomendasi, metode satu-keluar yang diuraikan di atas digunakan untuk setiap permintaan, dari kedua set data, di ketiga rekomendasi, dan jumlah rata-rata siklus dan item unik yang disajikan kepada pengguna lalu diukur.

Hasil untuk Travel dan PC dirangkum dalam Gambar 10 (a-d) sebagai grafik siklus rata-rata dan item untuk setiap algoritma dan grup queri. Hal ini terlihat jelas menunjukkan manfaat potensial dari strategi CP dan AS relatif terhadap *standard critiquing* (STD), terutama saat kesulitan maka queri meningkat. Dalam Perjalanan, menemukan bahwa untuk kueri sederhana dibutuhkan rata-rata 9,52 siklus rekomendasi (dan 28,55 item unik) untuk kritik standar untuk menemukan item target. Sebaliknya, CP mengambil 6,93 siklus (dan 15 item) dan AS mengambil 6,7 siklus (dan 14 item). Dengan kata lain, strategi AS mencapai pengurangan relatif dalam siklus 30% dan, dalam item, 50%, dibandingkan dengan STD. Karena berkurangnya kompleksitas domain PC (yaitu lebih sedikit fitur dan lebih sedikit kasus) manfaat di kueri sederhana kurang jelas daripada di Travel, dengan peningkatan kecil dalam jumlah item unik yang disajikan kepada

pengguna (8% dan 12% untuk AS) dan CP lebih dari STD), tetapi peningkatan kecil dalam jumlah siklus yang dibutuhkan.



Gambar. 10. Hasil efisiensi dari domain Travel dan PC.

Manfaat CP dan AS semakin ditingkatkan di kedua domain dengan pertanyaan yang semakin sulit. Dalam domain Perjalanan, jumlah siklus dan item yang diperlukan oleh STD meningkat menjadi 15,53 dan 46,59, untuk grup query moderat, dan dalam domain PC nilai yang sesuai adalah 5,81 dan 17,44. Namun, peningkatan siklus dan item berlangsung pada tingkat yang jauh lebih lambat untuk metode CP dan AS, yang mengarah ke peningkatan inkremental dalam manfaatnya relatif terhadap STD. Misalnya, dalam domain Perjalanan, untuk query yang sulit, 60% dramatis pengurangan jumlah item (dibandingkan dengan STD) dicapai untuk AS, dan pengurangan 52% untuk CP. Demikian pula, dalam domain PC, untuk query yang sulit, pengurangan item 44% dicapai untuk AS dengan pengurangan 31% untuk CP.

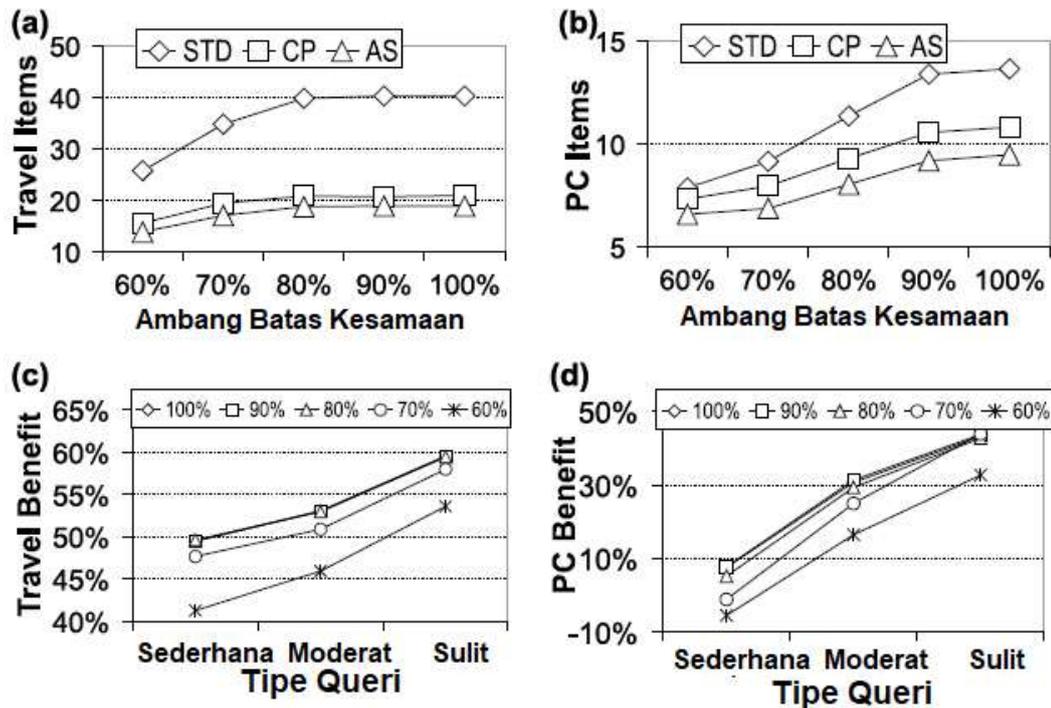
Hasil ini menunjukkan bahwa CP dan AS dapat memiliki dampak positif yang signifikan terhadap efisiensi pengkritik. Manfaat yang konsisten ditemukan untuk AS lebih dari CP dan manfaat CP dan AS terkait dengan meningkatnya kesulitan query - lebih sedikit siklus STD berarti lebih sedikit kesempatan untuk perbaikan oleh CP dan AS. Kekuatan hubungan ini dapat dinilai dengan mengukur korelasi antara siklus yang dibutuhkan oleh STD dan manfaat terkait yang dinikmati oleh CP dan AS (dalam hal siklus atau item). Korelasi ini ternyata melebihi 0,85 untuk berbagai kemungkinan kombinasi. Misalnya, menggabungkan siklus tolak ukur di kedua PC dan Perjalanan dan mengukur korelasi antara siklus ini dan manfaat terkait untuk AS, dalam hal pengurangan jumlah item, menghasilkan korelasi 0,87. Dalam praktiknya, hasil di atas menunjukkan bahwa pengurangan signifikan

dalam jumlah item bisa dicapai oleh AS untuk permintaan yang biasanya menuntut sekitar 3 atau lebih siklus, dengan pengurangan dalam siklus tersedia untuk permintaan yang membutuhkan sekitar 5 atau lebih siklus.

4.4 Toleransi Preferensi

Hasil di atas mengasumsikan bahwa dialog rekomendasi berakhir ketika item target yang telah ditentukan dipilih. Ini analog dengan pengguna yang mencari barang yang sangat spesifik tetapi, pada kenyataannya, pengguna cenderung lebih fleksibel dalam persyaratannya, menerima barang yang kurang optimal yang dekat dengan target. Percobaan diulangi, tetapi alih-alih mengakhiri dialog rekomendasi ketika target telah ditemukan, itu dihentikan setelah item ditemukan yang berada dalam kesamaan spesifik dari target. Selanjutnya diuji ambang kesamaan dari 60% hingga 100%; 100% sesuai dengan pengaturan sebelumnya di mana dialog berakhir dengan kasus target yang optimal. Hasilnya bisa dilihat pada Gambar 11 (a-d) dalam beberapa cara. Pertama, Gambar 11 (a & b) menyajikan hasil untuk Travel dan PC, grafik item rata-rata per dialog rekomendasi untuk ambang kesamaan yang berbeda. Dalam grafik ini, untuk alasan singkatnya, hasil untuk query moderat di setiap domain adalah ditunjukkan dengan hasil serupa yang diamati secara kualitatif untuk grup permintaan lainnya. Tidak mengherankan, karena kriteria keberhasilan menjadi lebih kaku, jumlah item unik yang disajikan kepada pengguna cenderung meningkat. Misalnya, dalam domain Perjalanan, jumlah item unik yang disajikan kepada pengguna oleh STD meningkat dari 25,76 (ambang 60%) menjadi sedikit di atas 40 (ambang batas 100%) dengan peningkatan yang sama untuk PC (7,8 item meningkat menjadi lebih dari 13 item). Sekali lagi melihat manfaat signifikan untuk CP dan AS, dengan AS secara konsisten mengalahkan CP. Misalnya, dalam Perjalanan, jumlah item yang diperlukan oleh AS pada ambang 60% adalah 13,92 (penurunan relatif terhadap STD 45,94%), yang meningkat menjadi 18,93 pada ambang 100% (pengurangan 53,01%). Manfaat AS (dan CP) terlihat meningkat dengan ambang kesamaan, setidaknya dalam kasus grup permintaan ini. Gambar 11 (c & d) mengonfirmasi ini di semua grup query dengan menggambar manfaat relatif (item) AS (relatif terhadap STD) terhadap kesulitan query. Dalam Perjalanan, manfaat meningkat dari minimum 41% hingga maksimum. 60%, dan dalam domain PC dari -5% hingga 44%; sekali lagi perlu dicatat bahwa -5% dikaitkan dengan jumlah rata-rata siklus 1,55 pada ambang kemiripan 60% dan merupakan satu-satunya situasi di mana dampak negatif ditemukan dalam jumlah metrik item untuk AS.

Diskusi Singkatnya, manfaat yang diamati sebelumnya untuk CP dan AS dijaga dalam kondisi penghentian yang lebih fleksibel. Memang manfaat AS meningkat ketika kondisi penghentian menjadi lebih kaku, hasil yang cenderung sangat penting dalam domain tempat pengguna cenderung mencari produk dan barang yang optimal atau hampir tidak optimal; misalnya, domain dengan produk mahal atau sangat terspesialisasi (misalnya. liburan, perhiasan, PC, peralatan audio visual kelas atas, dll.).



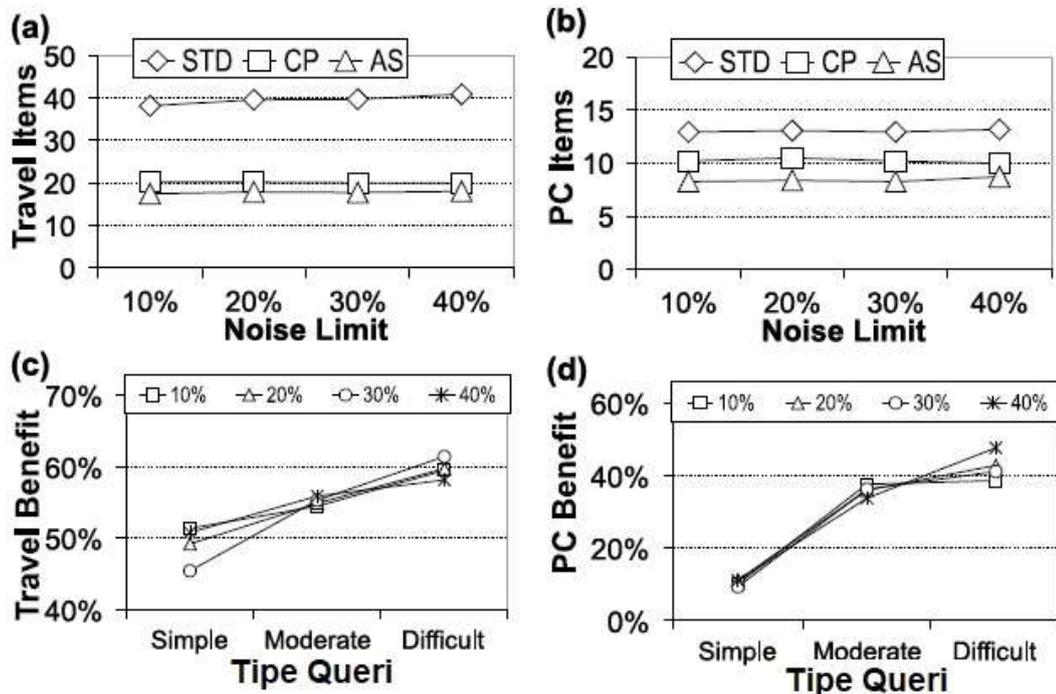
Gambar 11. Hasil toleransi preferensi dari domain Travel dan PC.

Perlu dicatat bahwa sebenarnya mungkin untuk mencapai rekomendasi optimal (ambang kemiripan 100%) menggunakan AS dengan jumlah item yang sama, atau lebih sedikit, daripada metode STD yang diperlukan untuk mencapai rekomendasi yang hanya 60% optimal, dalam Travel, atau 70% optimal, di PC. Misalnya, untuk grup permintaan moderat dalam Travel (lihat Gambar 11 (a)), AS memerlukan rata-rata 18,93 item pada ambang kemiripan 100%, sedangkan STD membutuhkan 25,76 item pada ambang kemiripan 60%. Hasil yang sebanding ditemukan, tetapi tidak dilaporkan di sini secara terperinci, untuk grup querri lainnya, meskipun mereka kurang jelas untuk domain PC.

4.6 Preference Noise

Sejauh ini mengasumsikan bahwa dalam setiap siklus pengguna mengkritik item yang paling mirip dengan target. Jika kita mematahkan asumsi ini, apa dampaknya terhadap manfaat CP dan AS yang diamati? Untuk menyelidiki masalah ini, mengulangi percobaan asli kecuali bahwa kebisingan dimasukkan ke dalam tahap mengkritik dengan mengganggu kesamaan, antara setiap kasus yang direkomendasikan dan target, dengan jumlah acak dalam batas kebisingan yang ditetapkan. Misalnya, batas noise 10% berarti bahwa setiap nilai kesamaan akan berubah dengan jumlah acak hingga +/- 10% dari nilai aktualnya. Ini berpotensi akan mengubah urutan item yang direkomendasikan selama setiap siklus sehingga, tergantung pada kedekatan nilai kesamaan asli dan ukuran batas kebisingan, item kedua atau bahkan ketiga yang paling mirip dengan target dapat

dikritik daripada yang paling barang serupa. Pendekatan ini meniru situasi di mana pengguna membuat preferensi kesalahan lebih sering ketika item yang direkomendasikan semua mirip dengan target ke tingkat yang kurang lebih setara. Angka Hasil 12 (a & b) grafik rata-rata jumlah item yang disajikan kepada pengguna versus batas kebisingan untuk queri moderat dalam domain Travel dan PC; hasil serupa diamati untuk pertanyaan sederhana dan sulit. Hasil ini agak mengejutkan dalam hal konsistensi mereka di berbagai level kebisingan (*noise*).



Gambar 12. Preferensi noise dihasilkan dari domain Travel dan PC.

Misalnya, tidak ada perubahan signifikan dalam jumlah rata-rata item yang ditemukan untuk STD, CP atau AS untuk tingkat kebisingan yang diperiksa, yang pada gilirannya menyebabkan manfaat yang konsisten diamati untuk AS dalam hal jumlah item yang disajikan kepada pengguna; lihat Gambar 12 (c & d). Misalnya, dalam domain PC menemukan bahwa manfaat AS hampir identik untuk berbagai tingkat kebisingan untuk kueri sederhana dan sedang, dan tidak berbeda secara signifikan bahkan untuk kueri yang sulit. Misalnya, di PC, manfaat AS untuk kueri sederhana secara konsisten sekitar 10% untuk semua tingkat kebisingan, naik menjadi sekitar 36% untuk kueri moderat, dan 41% untuk kueri yang sulit; pola yang hampir sama ditemukan di domain Perjalanan. Hasil serupa diamati untuk teknik CP dan sehubungan dengan jumlah siklus, tetapi tidak dilaporkan di sini secara rinci untuk alasan singkatnya. Diskusi Singkatnya, menemukan bahwa manfaat AS (dan CP) dipertahankan di seluruh tingkat kebisingan. Tetapi mungkin fitur yang paling menonjol dari hasil ini, adalah kurangnya sensitivitas yang diamati terhadap berbagai tingkat kebisingan. Dalam karya terbaru tentang murni umpan balik berbasis preferensi, ketika pengguna memilih item selain yang paling mirip dengan target, efisiensi rekomendasi menurun secara signifikan (L.

McGinty and B. Smyth, 2003). Tampaknya mengkritisi menawarkan perlindungan terhadap jenis kebisingan ini, mengompensasi preferensi yang kurang optimal. Tingkat perlindungan mungkin tergantung pada perbedaan dalam kesamaan dengan target antara item yang disukai / dikritik dan item lainnya dalam siklus rekomendasi. Dalam percobaan perbedaannya dalam kesamaan dengan target antara item terdekat dan terjauh dari target, dalam siklus rekomendasi tunggal, adalah sekitar 25-40% dari kesamaan item terdekat dengan target. Ini adalah perbedaan yang signifikan dan tidak diragukan lagi menjelaskan degradasi dalam efisiensi rekomendasi diamati ketika umpan balik berbasis preferensi digunakan sendiri. Namun, dari hasil menunjukkan bahwa mengkritisi memiliki kemampuan untuk mengatasi potensi penurunan kesamaan target untuk item yang disukai; mungkin karena penyaringan item oleh kritik yang dipilih membantu menjaga relevansi item yang tersisa untuk siklus berikutnya.

5. Kesimpulan

Umpan balik memainkan peran penting dalam banyak sistem pemberi rekomendasi yang dipersonalisasi dan berbagai jenis umpan balik memberikan keseimbangan yang berbeda sehubungan dengan biaya pengguna, ambiguitas umpan balik, keahlian pengguna, dan persyaratan antarmuka. Berorientasi pada konsumen merekomendasikan produk, sangat cocok untuk mengkritisi. Namun, dengan sendirinya bentuk umpan balik ini dapat menyebabkan dialog rekomendasi yang tidak efisien. Sebagai solusi, telah menjelaskan pemilihan adaptif, pendekatan baru terhadap item rekomendasi yang secara efektif menyeimbangkan pengaruh kesamaan dan keragaman selama pemilihan item. Seleksi adaptif telah terbukti secara signifikan meningkatkan efisiensi rekomendasi dalam berbagai kondisi eksperimental. Bahkan, manfaat AS meningkat dengan kesulitan query, dan pengurangan hingga 60% dalam jumlah item yang disajikan kepada pengguna (relatif terhadap kritik standar) telah diamati dengan andal yang disediakan sistem ini menjadi lebih baik.

Daftar Pustaka

- [1] D.W. Aha, L.A. Breslow, and H. Muoz-Avila. (2000). Conversational case-based reasoning. *Applied Intelligence*, 14:9–32.
- [2] D. Bridge. (2001). Product Recommendation Systems: A New Direction. In D. Aha and I. Watson, editors, *Workshop on CBR in Electronic Commerce at The International Conference on Case-Based Reasoning (ICCB-01)*, Vancouver, Canada.
- [3] R. Burke. (2000). Knowledge-based Recommender Systems. *Encyclopedia of Library and Information Systems*, 69(32).
- [4] R. Burke, K. Hammond, and B.C. Young. (1997). The FindMe Approach to Assisted Browsing. *Journal of IEEE Expert*, 12(4):32–40.
- [5] M. Doyle and P. Cunningham. (2000). A Dynamic Approach to Reducing Dialog in On-Line Decision Guides. In E. Blanzieri and L. Portinale, editors, *Proceedings of the Fifth European Workshop on Case-Based Reasoning*,

- (EWCBR-00), pages 49–60. Springer, Trento, Italy. Lecture Notes in Computer Science 19
- [6] M. Goker and C. Thompson. (2000). Personalized Conversational Case-based Recommendation. In E. Blanzieri and L. Portinale, editors, Proceedings of the Fifth European Workshop on Case-based Reasoning, (EWCBR-00), pages 99–111. Springer-Verlag,
 - [7] L. McGinty and B. Smyth. (2002). Comparison-Based Recommendation. In Susan Crow, editor, Proceedings of the Sixth European Conference on Case-Based Reasoning (ECCBR-02), pages 575–589. Springer, 2002. Aberdeen, Scotland.
 - [8] L. McGinty and B. Smyth. (2002). Deep Dialogue vs Casual Conversation in Recommender Systems. In F. Ricci and B. Smyth, editors, Proceedings of the Workshop on Personalization in eCommerce at the Second International Conference on Adaptive Hypermedia and Web-Based Systems (AH-02), pages 80–89. Springer, Universidad de Malaga, Malaga, Spain.
 - [9] L. McGinty and B. Smyth. (2003). The Role of Diversity in Conversational Systems. In D. Bridge and K. Ashley, editors, Proceedings of the Fifth International Conference on Case-Based Reasoning (ICCB-03). Springer, Troindheim, Norway.
 - [10] L. McGinty and B. Smyth. (2003). Tweaking Critiquing. In Proceedings of the Workshop on Personalization and Web Techniques at the International Joint Conference on Artificial Intelligence (IJCAI-03). Morgan-Kaufmann, Acapulco, Mexico.
 - [11] D. McSherry. (2001). Minimizing dialog length in interactive case-based reasoning. In Bernhard Nebel, editor, Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-01), pages 993–998. Morgan Kaufmann, Seattle, Washington.
 - [12] D. McSherry. (2002). Diversity-Conscious Retrieval. In Susan Crow, editor, Proceedings of the Sixth European Conference on Case-Based Reasoning (ECCBR-02), pages 219–233. Springer, Aberdeen, Scotland.
 - [13] J. Reilly and K. McCarthy and L. McGinty and B. Smyth. (2010). Dynamic Critiquing. Submitted to the European Conference on Case-Base Reasoning (ECCBR-04). To be held in Madrid, Spain.
 - [14] S. Rogers, P. Langley, B. Johnson, and A. Liu. (1997). Personalization of the automative information environment. In J. Herrmann R. Engels, B. Evans and F. Verdenius, editors, Proceedings of the workshop on Machine Learning in the real world; Methodological Aspects and Implications, pages 28–33, 1997. Nashville, TN.
 - [15] Shimazu. ExpertClerk. (2001). Navigating Shoppers' Buying Process with the Combination of Asking and Proposing. In Bernhard Nebel, editor, Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-01), pages 1443–1448. Morgan Kaufmann, Seattle, Washington, USA.
 - [16] H. Shimazu, A. Shibata, and K. Nihei. ExpertGuide. (2002): A conversational case-based reasoning tool for developing mentors in knowledge spaces. *Applied Intelligence*, 14(1):33–48.

- [17] B. Smyth and P. Cotter. (2000). A Personalized TV Listings Service for the Digital TV Age. *Journal of Knowledge-Based Systems*, 13(2-3):53–59.
- [18] B. Smyth and P. McClave. (2001). Similarity v's Diversity. In D. Aha and I. Watson, editors, *Proceedings of the International Conference on Case-Based Reasoning*, pages 347–361. Springer.
- [19] B. Smyth and L. McGinty. (2003). The Power of Suggestion. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-03)*. Morgan-Kaufmann, Acapulco, Mexico.