

SID



سرویس های
ویژه



سرویس ترجمه
تخصصی



کارگاه های
آموزشی



بلاگ
مرکز اطلاعات علمی



سامانه ویراستاری
STES



فیلم های
آموزشی

کارگاه های آموزشی مرکز اطلاعات علمی جهاد دانشگاهی

کارگاه آنلاین
بررسی مقابله ای متون (مقدماتی)

کارگاه آنلاین
پروپوزال نویسی و پایان نامه نویسی

کارگاه آنلاین آشنایی با پایگاه های اطلاعات علمی
بین المللی و
ترند های جستجو



خوشه بندی اسناد وب با استفاده از الگوریتم شبیه‌سازی تبرید ترکیبی بر اساس مدل رفتار کاربر

سعید مصطفی پور گندلی
دانشگاه آزاد اسلامی واحد آبادان
fdx_mg@yahoo.com

دکتر محمدحسین یکتایی
دانشگاه آزاد اسلامی واحد آبادان
my.yektaie@gmail.com

چکیده

در عصر حاضر مهم‌ترین منبع اطلاعاتی صفحات وبی است که بر روی اینترنت یافت می‌شوند که این صفحات به طرز فزاینده‌ای رو به افزایش هستند. از طرف دیگر کاربران برای رسیدن به صفحات هدف خود، مجبور هستند زمان زیادی را صرف جستجو در اینترنت کنند. پژوهشگران تلاش‌های زیادی در خصوص دسته‌بندی و خوشه‌بندی صفحات وب انجام داده‌اند. روش‌های موجود از خصوصیات ذاتی اسناد به‌منظور خوشه‌بندی استفاده می‌کنند. به نظر ما تعامل کاربران با سیستم‌های اطلاعاتی حاوی مطالب مفیدی است. به‌عنوان مثال اگر کاربران با سلاقی مشابه به دو سند دسترسی پیدا کنند، نشان‌دهنده‌ی مشابه بودن خود دو سند است. این مطالب می‌تواند کمک کنند تا اسناد براساس خصوصیات دیگری علاوه بر خصوصیات ذاتی آن‌ها (خصوصیات رفتاری کاربران) خوشه‌بندی شوند. در این پروژه تلاش در همین زمینه گردیده است و سعی شده با استفاده از استخراج اطلاعات رفتار کاربران، اطلاعات دقیق‌تر و بهتری با هزینه‌های محاسباتی کمتر از صفحات وب استخراج و در زمینه خوشه‌بندی اسناد استفاده شود؛ بنابراین در این تحقیق با استفاده از خوشه‌بندی و الگوریتم شبیه‌سازی تبرید ترکیبی تلاش شده تا هزینه‌های محاسباتی را کاهش و دقت روش خوشه‌بندی اسناد را افزایش یابد. دلیل اصلی این کار بهینه‌سازی روش خوشه‌بندی برای انتخاب مراکز بهینه برای خوشه‌ها است. نتایج این کار تحقیقاتی نشان می‌دهد بهینگی مناسب در زمینه دقت تشخیص در خوشه‌بندی صفحات وب و همین‌طور کاهش هزینه محاسباتی در خوشه‌بندی صفحات وب را نشان می‌دهد.

واژگان کلیدی: رفتار کاربر، خوشه بندی، الگوریتم شبیه سازی تبرید، شروع مجدد



مقدمه

اینترنت مملو از صفحات وب است. که آن را به سرعت به بزرگترین منبع اطلاعاتی که بشر تا کنون ساخته است بدل کرده است. رشد سرسام آور تولد وب سایت های جدید، باعث شده که جستجوی بهینه از نظر دقت در بازیابی صفحات در خواستی و زمان قابل قبول مورد توجه ویژه ای قرار گیرد. کاربران برای بدست آوردن اطلاعات مورد نظر خود با ابزاری به نام موتور های جستجوگر، که با رویکردی کاملاً منعطف در مقایسه با منابع چاپی عمل جستجو را انجام می دهند. منابع اطلاعاتی که در پس این صفحات نهفته است تلاش ما را برای دسترسی و یافتن اطلاعات مورد نظر دو چندان می کند. امروزه توسعه روز افزون منابع اطلاعاتی تا به آنجا رسیده که دیگر مشکل «یافتن اطلاعات» جدی تر از «نبودن اطلاعات» است. چراکه هم در زمینه کاهش زمان دسترسی کاربران به صفحات مورد نظرشان نقش کلیدی ایفا می کند؛ و هم از لحاظ نمایش محصولات و خدمات مرتبط با اطلاعات مورد درخواست کاربر مورد اهمیت است. در تحقیقات صورت گرفته، در زمینه داده کاوی (J. Furnkranz, 2010) اصولاً مدل رفتار کاربر برای خوشه بندی به کار گرفته نشده اند. بنابراین در این تحقیق تلاش خواهد شد؛ که برای خوشه بندی صفحات از مدل رفتار کاربر استفاده کرده و صفحات را با توجه به رفتار کاربران با استفاده از الگوریتم شبیه سازی تیرید، خوشه بندی شود.

روش تحقیق

در این تحقیق ما برای اولین بار برای حل مسائل خوشه بندی اسناد وب بر اساس مدل رفتار کاربر، از تکنیک شروع مجدد استفاده خواهیم کرد الگوریتم شبیه سازی تیرید را ارائه و نتایج مربوط را مورد تحلیل و بررسی قرار خواهیم داد (M. Eirinaki et al, 2010). اساس کار این تکنیک به این صورت است که تنوع را در الگوریتم بالا می برد و شرایط را برای رسیدن جواب های بهتر فراهم می سازد، یک دسته اصلی وجود دارد که تمامی اسناد موجود در آن قرار دارند این دسته متعلق به هیچ دسته علاقه مندی کاربران نیست و در واقع هنوز هیچ کاربری این صفحات را بازدید نکرده که نرخ دهی برای این صفحه را انجام دهد. کاربرانی که به سایت وارد می شوند، ابتدا تا زمانی که به نرخ چگالی براساس رابطه (۱) تعریف شده باشد، یعنی برای هر دسته تعداد مشخصی کاربر با نرخ های مورد نظر طراح وجود نداشته باشد، سؤالات ساخت پروفایل از آن ها پرسیده خواهد شد (R. O. Duda et al, 2002). سپس براساس تحلیل پاسخ ها دسته کاربر مشخص خواهد گردید. و در صورتی که دسته ها به اندازه مورد نظر طراح دارای نرخ و تعداد کاربران بود از کاربر سؤالات ساخت پروفایل پرسیده نمی شود و کاربر براساس بازدید صفحات و نرخ های رفتار که در ادامه توضیح داده می شود جزو دسته هم علاقه مندی قرار می گیرد (پس مسأله اولین کاربران با استفاده از سؤالات ساخت پروفایل و سپس تحلیل این پاسخ ها و ارائه به صورت بردار حل می گردد). (Q. He, 1999)

$$Class_i = \sum_{i=1}^n \frac{x_i}{y_i} \cdot N_i \quad \text{رابطه (۱)}$$

ابتدا تا زمانی که صفحاتی در دسته اصلی وجود دارند که توسط هیچ کاربری بازدید نشده اند عملیات خوشه بندی اسناد انجام نمی شوند در این حالت به تعداد کاربرانی که اسناد را بازدید کردند دسته خواهیم داشت در این حالت ممکن است تعدادی از صفحات فقط به وسیله یک کاربر و تعدادی دیگر به وسیله چند کاربر بازدید شده و همین باعث می شوند سند جزو چند دسته از کاربران قرار بگیرند، یعنی یک سند را چند کاربر دیده و نرخ های مختلفی به این سند براساس رفتارشان داده باشند برای دسته بندی سند ممکن است چندین کاربر یک سند را بازدید نمایند. پس ما در زیر رابطه ای ارائه می دهیم تا با



استفاده از این رابطه یک صفحه به دسته‌های کاربران مختلف بر اساس نسبت نرخ دهی رفتار کاربران متعلق گردد. برای این حالت ما رابطه‌ی (۲) تعریف نمودیم که به صورت زیر است:

$$P_j = \frac{x_i}{y_i} \quad \text{رابطه‌ی (۲)}$$

در رابطه (۲) برای هر سند P_j داریم نرخ داده شده کاربر $(x)_i$ به صفحه P تقسیم بر کل نرخ‌های داده شده کاربر i (که با Y نمایش داده‌ایم) برای سند P_j این رابطه را نسبت عضویت می‌نامیم. پس از نرخ دهی تمامی اسناد دسته اصلی با استفاده از رابطه نسبت عضویت (۲) داده‌های هر کاربر برای سند (نرخ دهی هر کاربر) و شماره کاربر را به صورت یک بردار تبدیل کرده و به رابطه‌هایی که در ادامه عنوان می‌کنیم برای خوشه‌بندی ارائه می‌دهیم. نرخ‌های داده شده هر کاربر و شماره سند را به عنوان ورودی الگوریتم خوشه‌بندی به خوشه‌بندی‌ها می‌دهیم این داده‌ها به صورت بردارهایی به صورت $i = \{(d_{i1}, \mu_i(d_{i1})), \dots, (d_{in}, \mu_i(d_{in}))\}$ به خوشه بندها ارسال می‌گردند در ادامه نحوه محاسبه $\mu_i(d_{i1})$ به عنوان درجه عضویت نرمال شده ارائه شده است. شبیه‌سازی تبرید یک الگوریتم جستجوی محلی یا یک روش فوق ابتکاری است که قادر است خود را از دام بهینه موضعی رها کند. راحتی در به کارگیری، همگرایی و استفاده از حرکات خاصی جهت دوری از قرارگیری در دام بهینه موضعی از جمله خصوصیات هستند که باعث شده‌اند تا این روش در دو دهه اخیر مورد توجه قرار گیرد (R. Baeza-Yates, 2004). نام گذاری این روش به نام شبیه‌سازی تبرید بیشتر به دلیل قیاس آن با فرآیند آنیل کردن فیزیکی در جامدات است که در آن یک ماده جامد تا یک دمای بالا حرارت داده شده و سپس به تدریج سرد می‌شود تا این که بتواند به پایدارترین شبکه کریستالی خود دست یابد (حداقل انرژی ممکن را داشته باشد). بدین ترتیب یک جامد با کیفیت بالا و عاری از نقص به دست می‌آید. شبیه‌سازی تبرید ارتباطی بین این رفتار ترمودینامیکی و فرآیند جستجو برای یافتن نقاط بهینه برقرار می‌کند. در SA، مسأله با یک حل اولیه شروع شده و با یک مجموعه از حرکات (با تغییر در مقادیر متغیرهای تصمیم) مطابق با طرح تبرید که شخص طراح از قبل تعریف کرده است، صورت می‌گیرد و زمانی که یا حل بهینه به دست آمد و یا مسأله در یک حل بهینه موضعی که دیگر امکان بهبود ندارد، بلوکه شد، متوقف می‌شود. الگوریتم برای آنکه در یک حل بهینه موضعی گرفتار نشود در فضای جواب به آرامی حرکت می‌کند (با توجه به مقدار هدف) این بهبود کنترل شده در مقدار هدف به وسیله قبول حرکات غیر بهینه با یک احتمال مشخص که با پیشرفت الگوریتم کاهش می‌یابد (مقدار آن هرچه الگوریتم به جلوتر می‌رود و همگام با کاهش دما، کاهش می‌یابد)، انجام می‌شود. این روش یک روش نسبتاً جدیدی است که برای حل بسیاری از مسائل ترکیبی مانند مسأله تخصیص مضاعف، فروشنده دوره گرد و غیره با موفقیت قابل قبولی بکار رفته است. حال به توضیح مختصر گام‌های الگوریتم شبیه‌سازی تبرید می‌پردازیم:

همسایگی: "همسایگی" به این صورت است که بعضی از مؤلفه‌ها به صورت کاملاً تصادفی تغییر می‌کنند. البته تعداد این گونه مؤلفه‌ها بسیار کم است اما در هر حال این تغییر تصادفی بسیار مهم است.
تابع برازندگی^۱: به عنوان معیاری برای ارزیابی هر جواب محاسبه می‌شود. نحوه محاسبه تابع برازندگی برای مسائل متفاوت فرق می‌کند.

مراحل الگوریتم SA به صورت زیر است:

S_0

یک حل اولیه را انتخاب کنید،

^۱ Neighborhood

^۲ Fitness Function

 t_0

یک دمای اولیه را انتخاب کنید

قدم‌های زیر را تکرار کنید:

$$S \in N(S_0)$$

یک حل جدید را به طور تصادفی انتخاب نمایید،

$$\delta = f(S) - f(S_0)$$

قرار دهید:

$$\delta < 0$$

اگر

$$S_0 = S$$

آنگاه قرار دهید:

در غیر این صورت یک عدد تصادفی بین (۱ و ۰) انتخاب نمایید.

$$S_0 = S$$

اگر $x < e^{-\delta/t}$ آنگاه قرار دهید:

تا زمانی که تعداد تکرارها برابر nrep شود.

$$t = \alpha(t)$$

قرار دهید:

تا زمانی که شرط توقف برقرار شود (رسیدن به دمای نهایی)

^۵ تقریبی از حل بهینه است.

بررسی روش‌های محاسبه شباهت بین خوشه‌ها

پس از اینکه ما میزان شباهت بین اسناد را به دست آوردیم اکنون می‌توانیم فاصله بین خوشه‌ها را هم به دست آوریم. از

سه معیار مختلف برای سنجش میزان شباهت خوشه‌ها استفاده می‌شود.

روش بهینه^۳: میزان شباهت بین دو خوشه برابر بیشترین شباهت بین دو سند مختلف از آن‌هاست، رابطه (۳).

رابطه (۳)

$$\text{Sim} = (C_i, C_j) = \max_{x \in C_i, y \in C_j} \text{Sim}(x, y)$$

روش بدبینانه^۴: میزان شباهت برابر با کمترین میزان شباهت بین دو سند مختلف است، رابطه (۴).

$$\text{Sim}(C_i, C_j) = \min_{x \in C_i, y \in C_j} \text{Sim}(x, y)$$

رابطه (۴)

روش مرکز خوشه^۵: میزان شباهت دو خوشه دقیقاً برابر میزان شباهت بین مرکز دسته‌های آن‌هاست.

$$\text{Sim}(C_i, C_j) = \text{Sim}(v_i, v_j)$$

رابطه (۵)

در رابطه (۵) به دست آوردن بازه تغییرات برای هر صفت در ویژگی‌های در نظر گرفته شده هر صفحه برای محاسبه میزان

تغییرات و برای نرمال‌سازی هر صفت در جستجوی، فضای جستجو برای هر صفت از رابطه زیر استفاده گردیده است.

³ optimal method⁴ pessimistic method⁵ cluster center method



$$X_{Center} = x_{min}(x_{max} - x_{min}) * rand$$

رابطه (۶)

در رابطه (۶) برای به دست آوردن نقاط تصادفی مرکز خوشه برای هر صفت است. چون الگوریتم فراابتکاری وظیفه پیدا کردن بهترین مراکز را برای هر صفت دارد این رابطه برای یافتن بهترین مقدار هر مرکز خوشه هر صفت استفاده گردیده است.

یافته ها :

۱- پیدا کردن پاسخ بهینه در فضای جستجو در زمان کمتر نسبت به رویکردهایی با بررسی تمامی داده‌ها

۲- کاهش هزینه محاسباتی برای یافتن پاسخ بهینه در فضای جستجو نسبت به رویکردهایی با بررسی تمامی داده‌ها

۳- افزایش دقت الگوریتم خوشه‌بندی با توجه به داده‌های پیداشده از طریق الگوریتم فرا ابتکاری

برای ساخت دیتاست این تحقیق از *log* فایل های ذخیره شده در پایگاه داده سایت <http://masjedsoleyman-samacollege.ir> استفاده شده است. سپس با استفاده از این *log* فایل ها داده‌های رفتاری کاربران برای خوشه‌بندی صفحات وب جداسازی و آماده گردید، جدول (۱).

جدول ۱: قسمتی از دیتاست به عنوان نمونه

User-ID	Action-Rating (print:10-save:6-rowse:3)
276725	3
276726	6
276727	3
276729	3
276729	6
276733	3
276736	10
276737	6
276744	6
276745	10
276746	3
276746	3



این صفحات شامل ۱۰۰۰ صفحه است که ۱۶۲ کاربر در یک دوره زمانی ۲ ماه بر اساس رفتارهایشان به آن‌ها نرخ داده‌اند سپس بر اساس محتوای صفحات به وسیله افراد خبره این صفحات به ۵ دسته کاربری هم علاقه دسته‌بندی گردیدند که تمام الگوریتم‌های خوشه‌بندی مقایسه شده شامل ۵ کلاس مختلف برای مقایسه با نتایج قضاوت افراد خبره انتخاب شده‌اند. زمان جستجو برای همه آزمایش‌ها یکسان در نظر گرفته می‌شود که برابر با ۴۵۰ ثانیه است.

زمان جستجو برای همه آزمایش‌ها یکسان در نظر گرفته می‌شود که برابر با ۴۵۰ ثانیه است. در ادامه اسامی نه الگوریتم به‌طور خلاصه آمده است:

SA: الگوریتم شبیه‌سازی تبرید

HSA: الگوریتم شبیه‌سازی تبرید ترکیبی

K-Means: الگوریتم K-Means

شاخص‌ترین و پرکاربردترین الگوریتم خوشه‌بندی مبتنی برافراز الگوریتم *K-Means* است، که در آن، هر خوشه با میانگین اشیای آن (مرکز خوشه) نمایش داده می‌شود (T. Kanungo et al, 2002). این الگوریتم هنگامی که خوشه‌ها به صورت ابرهای فشرده‌ی مجزا از هم هستند به خوبی کار می‌کند (E. Glover et al, 2002). این روش برای پایگاه‌های داده بزرگ نسبتاً کارا و ارتقا پذیر است، ولی اغلب به یک بهینه محلی منتهی می‌شود.

با توجه به محاسبات انجام‌شده، مقدار پارامترهای بهینه در همه الگوریتم‌ها تعیین می‌گردد که به شرح زیر است:

جدول ۲: پارامترهای بهینه الگوریتم SA

حداکثر تعداد تکرار در هر دما	دمای اولیه	نرخ کاهش دمای اولیه	عملگر همسایگی
۱۰۰	۲۰۰۰	۰/۹۵	دونقطه‌ای

جدول ۳: پارامترهای بهینه الگوریتم HSA

حداکثر تعداد تکرار در هر دما	دمای اولیه	نرخ کاهش دمای اولیه	عملگر همسایگی	حداکثر مقدار عدم بهبود
۱۰۰	۱۰۰۰	۰/۹	تک نقطه‌ای	۵۰

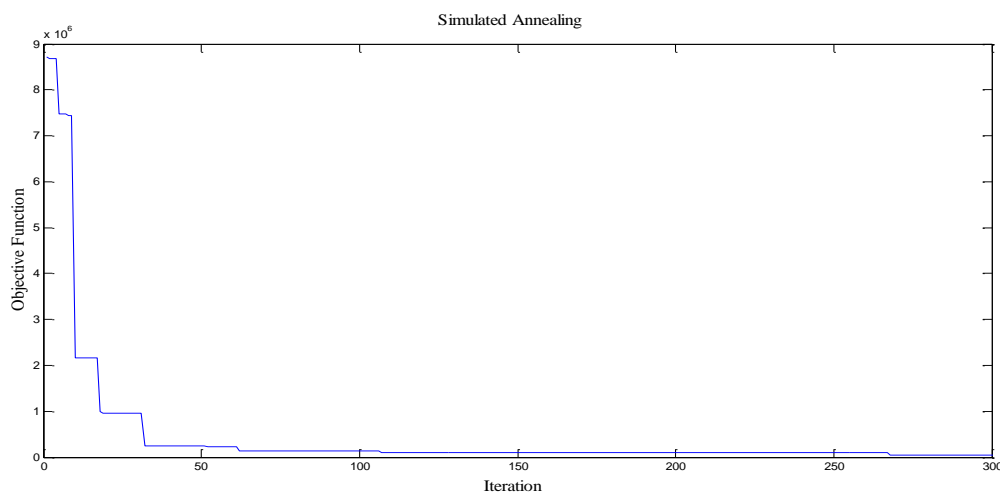


جدول ۴: مقادیر تابع هدف الگوریتم‌ها برای تابع هدف به ازای پارامترهای بهینه

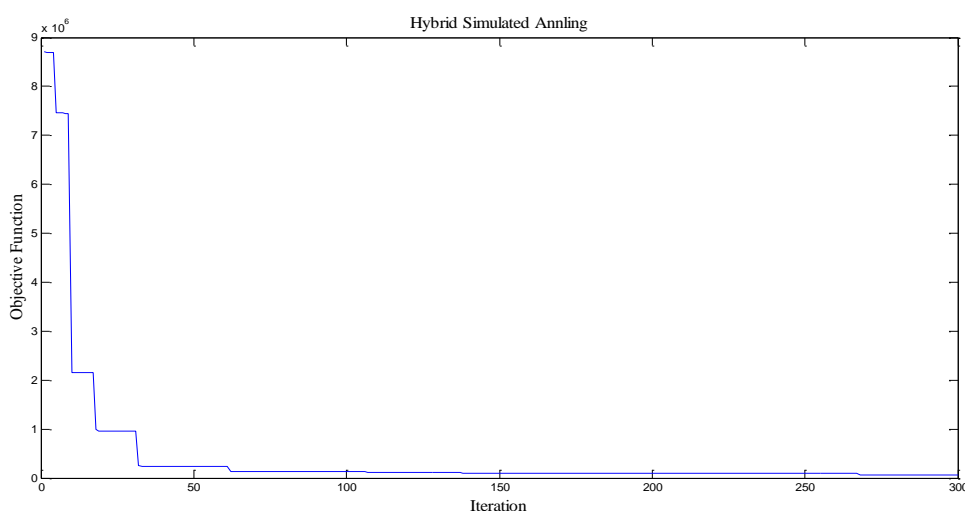
Sa	Has	K-Means	\bar{f}_k
50548.316	49974.082	49968.474	1
49968.486	49971.97	54739.18	2
49968.52	49974.758	125785.519	3
49968.509	49975.546	91097.523	4
49968.475	50018.387	49968.474	5
49968.477	49988.874	72200.224	6
49968.479	49975.375	86312.571	7
49968.483	49973.174	49968.474	8
49968.501	49991.217	55122.124	9
49968.478	50012.612	50548.24	10
50026.472	49985.6	68571.08	بهترین

نمودار همگرایی

در این قسمت برای الگوریتم‌های پیشنهادی نمودار همگرایی رسم می‌شود. بدین منظور در هر تکرار از الگوریتم، بهترین مقدار تابع هدف ارزیابی می‌شود. نمودارهای ۱ و ۲، نمودار همگرایی درازای یک‌زمان مشخص برای دو الگوریتم پیشنهادی است.



نمودار ۱: نمودار همگرایی برای الگوریتم SA



نمودار ۲: نمودار همگرایی برای الگوریتم HSA

بحث و نتیجه گیری

به همین دلیل همان طور که مشخص است هزینه جستجو در فضای رفتارهای کاربران بسیار کاهش پیدا کرده است؛ و به همین دلیل دقت خوشه بندی با استفاده از تکنیک جستجو تصادفی الگوریتم تبرید ترکیبی بسیار کاهش یافته است. زمانی که فضای عددی کوچک باشد این بهینگی کمتر مشخص خواهد بود، اما اگر فضای عددی در مسأله ای دارای تعداد زیادی باشد و در این فضا مقادیر دارای ترتیب خاصی نباشند، در این صورت میزان کاهش هزینه های محاسباتی برای پیدا کردن مقادیر بهینه، برای این فضای عددی بسیار بیشتر و چشم گیرتر از قبل خواهد بود. آنچه مسلم است روش های خوشه بندی دارای هزینه های کمتری در محاسبات و همین طور قابلیت انعطاف خوب در زمینه رویکردهایی بدون دانش اولیه در زمینه های گوناگون، همانند دستگاه های تشخیص تقلب و تشخیص نفوذ و یا خوشه بندی کاربران و یا صفحات و غیره می توان نام برد. مهم ترین دلیل انتخاب رویکردهای خوشه بندی برای مسائل جداسازی، انعطاف بالای این رویکرد است. چراکه آنچه این رویکردها برای جداسازی از آن بهره می برند متریک فاصله است؛ که این قابلیت به مسائل گوناگون این امکان را می دهد که با استفاده از تبدیل تابع هدف مسأله به فاصله بتواند تقریباً هرگونه مسأله جداسازی را حل نماید. از طرف دیگر دانش اولیه در بسیاری از مسائل دارای اهمیت فراوانی است و به وجود آوردن این دانش برای روش های دسته بندی یا عملاً غیرممکن است یا اینکه هزینه های بسیار زیادی از جمله هزینه های زمانی و مالی و تخصصی دارد. برای همین در روش هایی که عملاً ایجاد دانش



اولیه غیرممکن است، باعث می‌شود تنها رویکرد قابل انجام، رویکردهای بر مبنای خوشه‌بندی و روش‌های مرتبط با این زمینه باشد؛ اما مسأله‌ی مهم در این رویکردها، گیر افتادن هر بهینه محلی و بالا بودن هزینه محاسباتی برای انتخاب مرکز بهینه هر خوشه است؛ و در ضمن آنچه می‌تواند دقت این رویکردها را به میزان چشمگیری افزایش دهد، استفاده از روش‌های جدیدی و هوشمندانه یا تکاملی برای پیدا کردن مراکز بهینه با هزینه کم محاسباتی و زمانی است. در این تحقیق برای اولین بار از الگوریتم‌های شبیه‌سازی تبرید ترکیبی، جهت حل مسأله، خوشه‌بندی اسناد وب براساس مدل رفتار کاربر استفاده شده است. همچنین در این مطالعه برای نخستین بار جهت اجرای هرچه بهتر الگوریتم‌ها از مکانیسم شروع مجدد استفاده شد. در ادامه تحقیق مقایسه‌ای جامع بین انواع پارامترها و عملگرهای مربوط به هر یک از الگوریتم‌های ارائه‌شده، جهت یافتن و تنظیم بهترین عملکرد برای هر الگوریتم صورت پذیرفت. بعد از یافتن بهترین پارامترها و عملگرها برای اجرای بهتر هر الگوریتم، روش‌های فراابتکاری توسعه داده‌شده با یکدیگر به رقابت پرداختند که نتایج محاسباتی و تحلیل نتایج حاکی از این بود الگوریتم شبیه‌سازی تبرید ترکیبی دارای عملکرد بهتری است. در مقایسه کلی و جامع در بین سه الگوریتم ارائه‌شده، مشاهده می‌شود که الگوریتم شبیه‌سازی تبرید ترکیبی با میانگین تابع هدف ۴۴۹۹۸۵٫۶ دارای بهترین عملکرد است.

منابع

- [1] M. Eirinaki, M. Vazirgiannis, I. Varlamis, SEWeP: using site semantics and a taxonomy to enhance the Web personalization process, Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, 2010.
- [2] R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification And Scene Analysis, John Wiley & Sons, 2000R. Web, Statistical Pattern Recognition, John Wiley & Sons, 2002.
- [3] J. Furnkranz. Web mining. The Data Mining and Knowledge Discovery Handbook, pages. Springer, 2010, 899- 920.
- [4] R. Baeza-Yates. Web mining in search engines. Proceedings of the 27th conference on Australasian computer science, Volume 26, 2004, 3-4.
- [5] Q. He, A Review of Clustering Algorithms as Applied in IR, Graduate School of Library and Information Science University of Illinois at Urbana-Champaign, 1999
- [6] T. Kanungo, Mount D. M., Netanyahu N., Piatko C., Silverman R. and Wu A. Y., 2002. A local search approximation algorithm for K-Means clustering. Computational Geometry: Theory and Applications, SoCG'02, pp. 89-112.
- [7] E. Glover, K. Tsioutsoulklis, S. Lawrence, D. Pennock, G. Flake. Using web structure for classifying and describing web pages. In International World Wide Web Conference, May 2002.

SID



سرویس های ویژه



سرویس ترجمه تخصصی



کارگاه های آموزشی



بلاگ مرکز اطلاعات علمی



سامانه ویراستاری STES



فیلم های آموزشی

کارگاه های آموزشی مرکز اطلاعات علمی جهاد دانشگاهی

توجه: بررسی مقاله ای متون (مقدماتی)

کارگاه آنلاین
بررسی مقابله ای متون (مقدماتی)

PROPOSAL
پروپوزال

توجه: پروپوزال نویسی و پایان نامه نویسی

کارگاه آنلاین
پروپوزال نویسی و پایان نامه نویسی

ISI
Scopus

توجه: آشنایی با پایگاه های اطلاعات علمی بین المللی و ترند های جستجو

کارگاه آنلاین آشنایی با پایگاه های اطلاعات علمی بین المللی و ترند های جستجو