

چکیده

استفاده همزمان از اطلاعات ساختاری و اطلاعات پیمایش کاربران یکی از چالش‌های مطرح در بهبود کارایی الگوریتم‌های شخصی‌سازی وب می‌باشد. در این مقاله پس از معرفی معیار وزن دهی، الگوریتمی ترکیبی که از اطلاعات پیمایش کاربران و پیوند بین صفحات به منظور پیشنهاد صفحات به کاربران استفاده می‌کند، ارائه شده است. معیار معرفی شده برای محاسبه وزن صفحات مشاهده شده توسط کاربران از "مدت زمان مشاهده صفحه" و "فرکانس مشاهده صفحه" استفاده می‌کند که به خوبی میزان اهمیت و علاقه کاربران به آن صفحه را نشان می‌دهد. الگوریتم ارائه شده دو مشکل اساسی را در شخصی‌سازی صفحات وب حل می‌کند. مشکل اول پیشنهاد صفحات جدیدی است که اخیراً به سایت اضافه شده اند و مشکل دوم کاهش دقت الگوریتمها با افزایش تعداد صفحات پیشنهادی می‌باشد. در الگوریتم ارائه شده اولین صفحه با استفاده از قوانین انجمنی وزن دار جدید ارائه شده پیشنهاد می‌شود. سپس این صفحه با استفاده از الگوریتم HITS و صفحاتی که با آن در یک دسته بندی هستند بسط داده می‌شود تا صفحاتی که اخیراً به سایت اضافه شده اند نیز فرصت حضور در مجموعه صفحات پیشنهادی را داشته باشند. برای دسته بندی صفحات الگوریتمی بر اساس آتاماتای یادگیر و الگوریتم های افزاز گراف ارائه شده است. نتایج شبیه‌سازی الگوریتم در داده‌های واقعی نشان داده است که کارایی الگوریتم پیشنهادی بالا می‌باشد و دانش بدست آمده از سیستم مذکور به طور قابل ملاحظه ای کیفیت پیشنهادات را بهبود داده است و مشکلات ذکر شده را در حد قابل توجهی کاهش داده است.

کلمات کلیدی

داده کاوی وب، قواعد انجمنی وزن دار، اتوماتای یادگیر، الگوریتم HITS

^{۹۲} دانشکده مهندسی برق، رایانه و فناوری اطلاعات دانشگاه آزاد اسلامی قزوین ایران forsati@mrl.ir

^{۹۳} دانشکده مهندسی کامپیوتر و فناوری اطلاعات دانشگاه صنعتی امیرکبیر تهران ایران mmeybodi@aut.ac.ir