

شخصی سازی صفحات وب با استفاده از اتوماتای یادگیر توزیع شده

رعنا فرصتی^۱، محمد رضا میبیدی^۲ و مهرداد مهدوی^۳

(۱) دانشکده مهندسی برق، رایانه و فناوری اطلاعات، دانشگاه آزاد اسلامی، قزوین، ایران

(۲) دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران، ایران

(۳) دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف، تهران، ایران

forsati@mrl.ir, mmeybodi@aut.ac.ir, mahdavi@ce.sharif.edu

چکیده - استفاده همزمان از اطلاعات ساختاری و اطلاعات پیمایش کاربران یکی از چالش‌های مطرح در بهبود کارایی الگوریتم‌های شخصی سازی وب می‌باشد. در این مقاله الگوریتمی ترکیبی مبتنی بر اتوماتای یادگیر توزیع شده و الگوریتم *PageRank* ارائه شده است. الگوریتم پیشنهادی از اطلاعات پیمایش کاربران و پیوند بین صفحات به منظور پیشنهاد صفحات به کاربران استفاده می‌کند. الگوریتم ارائه شده را همچنین می‌توان برای تغییر پیوند صفحات به منظور هدایت بهتر کاربران استفاده کرد. بر خلاف الگوریتم‌های شخصی سازی موجود که تنها از اطلاعات پیمایش کاربران استفاده می‌کنند، الگوریتم ارائه شده اولین روش گزارش شده مبتنی بر اتوماتای توزیع شده می‌باشد که همزمان از اطلاعات پیمایش کاربران و پیوند بین صفحات برای پیشنهاد صفحات استفاده می‌کند. در الگوریتم ارائه شده یک اتوماتای یادگیر به هر صفحه وب تخصیص داده می‌شود. هر اتوماتای یادگیر، بر اساس اطلاعات پیمایش کاربران احتمال گذار بین صفحات را یاد می‌گیرد. بر اساس احتمالات گذار و اهمیت هر صفحه که با استفاده از الگوریتم *PageRank* محاسبه می‌شود، عملیات شخصی سازی انجام می‌شود. بر خلاف الگوریتم *PageRank* موجود که اهمیت هر صفحه بر اساس ساختار پیوندی صفحات محاسبه می‌شود الگوریتم ارائه شده همزمان از اطلاعات ساختار پیوندی صفحات و پیمایش کاربران برای محاسبه اهمیت صفحات استفاده می‌کند. نتایج شبیه‌سازی الگوریتم در داده‌های واقعی نشان داده است که کارایی الگوریتم پیشنهادی به ۹۰٪ می‌رسد در حالیکه پیچیدگی زمانی اجرای آن نیز پایین می‌باشد.

کلید واژه- اتوماتای یادگیر، الگوریتم *PageRank*، داده‌کاوی استفاده از وب.

۱- مقدمه

شخصی کردن وب به یک پدیده محبوب به منظور سفارشی کردن محیط‌های وب تبدیل شده است. هدف از سیستم‌های شخصی‌ساز فراهم کردن نیازهای کاربران، بدون اینکه به طور صریح آن‌ها را بیان کنند یا نشان بدهند، می‌باشد [۳]. شخصی سازی وب مجموعه‌ای از عملیات است که تجربه وب را برای یک کاربر خاص یا مجموعه‌ای از کاربران سازمان دهی می‌کند. شخصی سازی وب می‌تواند به عنوان فرآیندی برای سفارشی کردن محتوا و ساختار وب سایت بر حسب نیازهای شخصی و ویژه هر کاربر باشد [۴]. سایت از طریق پررنگ کردن لینک‌های موجود، ایجاد پویای

وب طی یک فرآیند آشفته و غیر متمرکز رشد می‌کند و این روند منجر به تولید حجم وسیعی از مستندات متصل به یکدیگر گشته است که از هیچ گونه سازماندهی منطقی برخوردار نیستند. در حال حاضر موتور جستجوی Google بیش از ۳ بلیون صفحه وب را شاخص گذاری کرده است. بنابراین وب به مجموعه بزرگی از داده‌های ساخت یافته و نیمه ساخت یافته تبدیل شده است که کاربران آن از همپوشانی داده‌ها رنج می‌برند. برای حل این مشکل،