

خوشه بندی الگوهای دسترسی وب با استفاده از اتوماتای یادگیر و منطق فازی

زهره اناری^۱؛ محمد رضا میبیدی^۲؛ بابک اناری^۳

چکیده

علاقه کاربران وب می تواند توسط صفحات وب ملاقات شده و مدت زمان بروی این صفحات در طی گشت و گذار آنها دروب مشخص شود. پارامتر مدت زمان یک صفحه وب که در لاگ فایلها ذخیره می شود، پارامتر مهمی در آنالیز رفتار حرکتی کاربران وب به حساب می آید. از آنجائیکه زمان به صورت عددی بیان می شود، می توان از مفاهیم فازی برای پردازش آن و ایجاد متغیرهای زبانی استفاده نمود. در این مقاله یک الگوریتم دو مرحله ای برای خوشه بندی الگوهای دسترسی وب با استفاده از ترکیب اتوماتای یادگیر و منطق فازی پیشنهاد می کنیم. در اولین مرحله هر الگوی دسترسی وب از لاکهای وب به الگوی دسترسی فازی وب تبدیل می شود، که یک بردار فازی، متشکل از متغیرهای زبانی فازی یا صفر می باشد. هر عنصر در الگوهای دسترسی فازی وب نشان دهنده صفحه وب ملاقات شده و مدت زمان بر روی این صفحه وب می باشد. سپس با استفاده از اتوماتای یادگیر هر الگوی دسترسی فازی وب را در نزدیکترین خوشه مربوط قرار می دهیم، با این کار یک خوشه بندی اولیه بر روی الگوهای دسترسی وب انجام می گیرد، همچنین مراکز اولیه خوشه ها نیز تعیین می شود. در دومین مرحله این خوشه های اولیه که هر کدام دارای صفر یا چند الگوی دسترسی وب هستند توسط الگوریتم خوشه بندی وزندار **weighted fuzzy c-means** مورد استفاده قرار گرفته و برحسب وزن هر خوشه که از روی تعداد الگوهای دسترسی قرار گرفته در هر خوشه بدست آمده اند و همچنین مراکز خوشه ها که توسط اتوماتای یادگیر تعیین شده است، خوشه بندی مجدد می شوند. با این کار خوشه های نهایی از روی خوشه های اولیه بدست می آیند. نتایج آزمایشها که بر روی چند لاگ داده واقعی وب تست شده است، کارایی بالای الگوریتم پیشنهادی را در مقایسه با سایر روشهای موجود نشان می دهد.

کلمات کلیدی

اتوماتای یادگیر، خوشه بندی، الگوهای دسترسی وب، متغیر فازی

۱. زهره اناری، دانشجوی کارشناسی ارشد کامپیوتر، گرایش نرم افزار، دانشگاه آزاد اسلامی واحد

شبهستر، zanari322@yahoo.com

۲. دکتر محمد رضا میبیدی، استاد و عضو هیئت علمی دانشگاه، دانشگاه صنعتی امیر

کبیر، mmeybodi@aut.ac.ir

۳. بابک اناری، عضو هیئت علمی دانشگاه، دانشگاه آزاد اسلامی واحد

شبهستر، anari322@yahoo.com