



کنگره بین المللی علوم و مهندسی

آلمان – هامبورگ

اسفند ماه ۱۳۹۶

ارائه روش جدید برای وب کاوی با استفاده از الگوریتم چند معیاره

فازی

مجید احمدی^{۱*}، سمیه حسن پور^۲، کیومرث عبدی^۳

۱- دانشگاه شهید مدنی آذربایجان، تبریز، ایران ، M.ahmadi@azaruniv.edu

۲- دانشگاه فنی و حرفه ای دخترانه الزهرا(س)، تبریز، ایران ، somayeh_hasanpoor@yahoo.com

۳- دانشگاه فنی و حرفه ای دخترانه الزهرا(س)، تبریز، ایران ، kiomarsabdi@yahoo.com

چکیده

وب کاوی و داده کاوی ارتباط بسیار نزدیکی با یکدیگر دارند. داده کاوی فرآیند ارائه پرس و جوها و استخراج الگوها و اطلاعات مفید و ناشناخته از داده هایی است که معمولاً در پایگاه داده ها ذخیره شده اند. در واقع بسیاری از روش های داده کاوی قابل استفاده در وب کاوی هستند. اما گستره وب کاوی وسیع تر از داده کاوی است. یکی از مهمترین چالش های دنیای امروز در زمینه وب کاوی، قرار دادن داده ها و صفحات وب مشابه در یک گروه است. مفهوم داده های مشابه از دیدگاه خوشه بندی متفاوت بوده و می تواند چندین پارامتر مانند مشابهت متنی، ساختاری، کاربردی، تعداد کلمات استفاده شده، کلمات هم معنی و مترادف و ... را شامل شود و باید صفحات و داده ها طوری خوشه بندی شود که در هنگام جستجو، صفحات و داده های مرتبط ارائه گردد و همچنین طبقه بندی و نمایه سازی صفحات و داده ها مناسب واقع گردد. در این مقاله، روش جدیدی مبتنی بر فازی با هدف افزایش دقت جستجوی موضوعات مرتبط با مورد جستجو ارائه شده است که صفحات و داده های مشابه را خوشه بندی درست نموده و با استفاده از وزن و میزان ارجحیت هر یک از پارامترهای همگونی، محاسبه تشابه داده ها و صفحات انجام یافته است. الگوریتم ارائه شده در نرم افزار متلب شبیه سازی شده و نتایج با الگوریتم جدید ارائه شده مورد مقایسه و ارزیابی قرار گرفته است.

واژه های کلیدی: وب کاوی، داده کاوی، خوشه بندی وب، تشابه صفحات

۱- مقدمه (B Nazanin 14pt پررنگ)

با توسعه سیستم های اطلاعاتی، داده به یکی از منابع پراهمیت سازمان ها مبدل گشته است. بنابراین روش ها و تکنیک هایی برای دستیابی کارا به داده، اشتراک داده، استخراج اطلاعات از داده و استفاده از این اطلاعات، مورد نیاز می باشد. با ایجاد و گسترش وب و افزایش چشمگیر حجم اطلاعات، نیاز به این روش ها و تکنیک ها بیش از پیش احساس می شود. وب، محیطی وسیع، متنوع و پویا است