

تحلیل رفتار کاربران فارسی زبان در موتور جستجوی بومی

محمد صادق زاهدی^۱، معصومه عظیم زاده^۱، نوید فرهادی^۱، علی محمد زارع بیدکی^۲

^۱ گروه سکویای فناوری اطلاعات و فضای مجازی، پژوهشگاه ارتباطات و فناوری اطلاعات، مرکز تحقیقات مخابرات، تهران

{zahedi, azim_ma, farhadi}@itrc.ac.ir

^۲ دانشکده مهندسی برق و کامپیوتر، دانشگاه یزد، یزد

alizareh@yazduni.ac.ir

چکیده

امروزه با رشد روز افزون اطلاعات موجود در وب، موتورهای جستجو تبدیل به ابزاری مهم و محبوب برای یافتن اطلاعات مورد نیاز کاربران در وب شده است ولی با این وجود هنوز پژوهشی در زمینه تحلیل رفتار کاربران ایرانی در موتورهای جستجوی بومی صورت نگرفته است لذا در این مقاله، سعی شده رفتار کاربران ایرانی که از موتور جستجوی پارسی جو استفاده نموده‌اند را به کمک تحلیل فایل ثبت وقایع (log) این موتور جستجو بررسی نماییم. تحلیل‌های صورت گرفته در سطوح مختلف پرس‌وجو، نشست، کلیک کاربر و وبسایت‌های پربازدید در بازه زمانی حدود دو ماه صورت گرفته است. نتایج کسب شده حاکی از آن است که کاربران ایرانی بیشتر در مورد موضوعات کامپیوتر و اینترنت، افراد و مکان‌ها، اخبار، علمی و اجتماعی جستجو می‌نمایند و نیازهای اطلاعاتی خود را در قالب نشست‌های کوتاه مدت و با تعداد کمی از پرس‌جوها در هر نشست مطرح نموده‌اند. علاوه بر آن میانگین تعداد کلیک کاربران برای جستجوی عکس بیشتر از جستجوی وب بوده است و میانگین طول پرس‌و‌جوهای کاربران نیز حدود ۲۰۷۸ بوده است.

کلمات کلیدی

تحلیل فایل ثبت وقایع، موتور جستجو، تحلیل رفتار کاربران، نشست، کلیک کاربران.

۱- مقدمه

تحلیل لاگ موتور جستجو است که بیشتر به صورت کمی می‌باشد [۵، ۶]. به دلیل اینکه فایل لاگ حاوی تعداد زیادی تعاملات بین کاربر و موتور جستجو می‌باشد و همچنین کاربر از اینکه مورد تحلیل قرار گرفته، آگاهی ندارد این روش مناسب‌تر است. لذا در این مقاله ما از روش تحلیل لاگ استفاده نموده‌ایم.

در حالت کلی یک فایل لاگ موتور جستجو در واقع تعامل بین کاربر و موتور جستجو را ثبت می‌کند و شامل دو مولفه اصلی است. این دو مولفه عبارتند از ورودی کاربر به موتور جستجو و دیگری خروجی ارائه شده به کاربر توسط موتور جستجو، که از این موارد برای تجزیه و تحلیل رفتار کاربران موتور جستجو استفاده می‌کنند.

تحقیقات قبلی در زمینه تحلیل لاگ موتور جستجو نشان داده است که کاربران نقاط مختلف جهان با توجه به زبان و فرهنگ خاص آن کشور رفتار متفاوتی در مواجهه با موتور جستجو دارند. به عنوان مثال میانگین طول پرس‌و‌جوهای کاربران کره‌ای کمتر از کاربران اروپایی و آمریکایی می‌باشد [۷] و یا کاربران آمریکایی بیشتر در مورد تجارت الکترونیک جستجو نموده‌اند در حالی که تمرکز کاربران اروپایی بیشتر در مورد "افراد و مکانهای مختلف" بوده است [۱].

موتور جستجوی پارسی جو یکی از موتور جستجوی بومی برای زبان فارسی می‌باشد که شامل سرویس‌های جستجوی وب، عکس،

اهمیت موتورهای جستجو به عنوان دروازه ورودی اینترنت بسیار مهم است زیرا این موتورهای جستجو هستند که کاربران به وسیله آن‌ها در اینترنت گشت و گذار نموده و حتی گاهی اوقات، مخاطب را به سمت و سویی هدایت می‌کنند. بر اساس [۱] موتور جستجو درگاه ورود ۸۰ درصد از کاربران وب برای یافتن اطلاعات موردنیاز خود می‌باشد. این اهمیت موضوع تحلیل رفتار کاربران موتور جستجو را مشخص می‌کند، اینکه کاربران چگونه موتورهای جستجو را برای یافتن اطلاعات مورد نیاز خود به کار می‌گیرند. در ایران نیز موتورهای جستجوی یکی از پراستفاده‌ترین کاربردها در وب می‌باشند. بنا بر آمار سایت الکسا، از بین ۵ سایت پربازدید ایران، ۳ سایت مربوط به موتورهای جستجو بوده است.

تحلیل رفتار کاربران موتورهای جستجو به دو صورت امکان پذیر است: یکی از طریق ارزیابی‌های انسانی و تحلیل مشاهدات کاربران [۳]، [۴] که این روش بیشتر به صورت کیفی می‌باشد و به دلیل محدودیتی که در تعداد افراد شرکت کننده در آزمایش و همچنین هزینه زیادی که برای این کار وجود دارد، لذا انجام این آزمایش در مقیاس بالا با دشواری و محدودیت‌های خاصی مواجه است. روش دوم مبتنی بر