

عنوان مقاله:

بررسی و کشف رونوشت های نزدیک انجمن های وب

محل انتشار:

همایش ملی مهندسی کامپیوتر و فناوری اطلاعات (سال: 1392)

تعداد صفحات اصل مقاله: 8

نویسندگان:

دکتر حسن نادری - استادیار دانشگاه علم و صنعت

فریبا سلیمانی

خلاصه مقاله:

فن آوری های انجمن جستجوی فعلی قادر به شناسایی دنباله ها در مندرجات نزدیک به رونوشت (تکرار) و گروه بندی این دنباله ها در نتایج جستجو نمیباشند. در نتیجه کاربران انجمن با نتایج جستجوی رونوشت شده بار زیادی بر دوش دارند که ترجیح می دهند دنباله های جدید بدون سعی در یافتن آنچه موجود است، ایجاد کنند. بنابراین در این مقاله ما دلایل معمول را که منجر به رونوشت و ایجاد الگوریتم کشف رونوشت جدید برای دنباله های انجمن میشود را شناسایی می کنیم. این الگوریتم با استفاده از یک پژوهش بزرگ از انجمن در دنیای واقعی که بیش از یک میلیون کاربر را به خدمت گرفته است، کامل می شود. ما این کار را با الگوریتم های فعلی مشابه برای کشف رونوشت ها در صفحات وب مورد مقایسه قرار می دهیم. نتایج اولیه ی ما نشان میدهد که ما به طور چشمگیری این الگوریتم ها را به کار گرفته ایم و ما موفق به گروه بندی دنباله های انجمن با دقت 74 درصد شده ایم.

کلمات کلیدی:

کشف رونوشت نزدیک، انگشت نگاری، گروه بندی دنباله

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/254202>

