

ردیابی چرخه عمر موضوعات خبری در شبکه های اجتماعی

سمیه افخمی اردکانی^۱، محمدرضا ملاخلیلی^۲

^۱ دانشجوی کارشناسی ارشد، دانشگاه آزاد اسلامی واحد میبد، یزد، ایران

Afkhami_۱۳۸۵@yahoo.com

^۲ استادیار، دانشگاه آزاد اسلامی واحد میبد، یزد، ایران

Mollakhalili@maybodiau.ac.ir

چکیده

با توسعه سیستم های اطلاعاتی، داده به یکی از منابع پراهمیت مبدل گشته است. کاوش در داده های کیفی و در این جا متن های خبری جهت یافتن دانش کاربردی موضوع نسبتا جدید و جذابی است. این کار به ویژه در رده های خبری مورد توجه پژوهش گران زیادی قرار گرفته است. هدف این مقاله تعیین چرخه ی عمر مقاله های خبری برخط با استفاده از واکنش شبکه های اجتماعی است. در این تحقیق نشان داده می شود که می توان از واکنش شبکه های اجتماعی برخط برای پیش بینی الگوی باز دیده ها استفاده کرد. و با ترکیب باز دیده های وب سایت و واکنش شبکه های اجتماعی طول عمر رده های خبری را به دست آورد. برای این منظور مجموعه داده ی یک ماهه فایل ثبت وقایع سایت خبری ایران نیدز، با بیش از ۱۲۰۰۰۰۰ رکورد مورد بررسی قرار می گیرد. از روش های رگرسیون و دسته بندی و با استفاده از روش های آماری جهت ارزیابی طول عمر خبر استفاده می گردد. نتایج حاصل از تحقیق نشان می دهد که طول عمر رده خبری سرگرمی و اجتماعی به طور متوسط بیش از ۲۰ روز می باشد.

کلمات کلیدی

طول چرخه عمر، شبکه های اجتماعی، داده کاوی

۱- مقدمه

شبکه های اجتماعی به دلیل داشتن خصوصیات «سیال بودن»، «جهانی بودن»، «فرازمانی»، «فرامکانی» فضای مناسبی را برای بیان آراء و عقاید در اختیار کاربران قرار می دهند. البته امکان گذاشتن اظهار نظر^۱ بر روی مطالب این شبکه ها توسط دوستان، اهمیت شبکه های اجتماعی مجازی را دو چندان می کند [۲]. هدف این مقاله به تعیین چرخه ی عمر مقاله های خبری آنلاین با استفاده از واکنش شبکه های اجتماعی پرداخته شده است. مفهوم چرخه عمر در دهه های اخیر واژه ای شناخته شده است. بر اساس این مفهوم تمامی موجودات همگی از منحنی عمر یا چرخه ی عمر پیروی می کنند. اینگونه موجودات متولد می شوند، رشد می کنند، به پیری میرسند و در نهایت می میرند. تئوری چرخه ی عمر چنین فرض می کند که اخبار هم مانند تمامی موجودات

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 که کاربران متعدد اسناد خود را در آن منتشر می کنند. در حال حاضر بیش از دو بیلیون صفحه در وب موجود است و این تعداد با نرخ ۷.۳ میلیون صفحه در روز افزایش می یابد. با توجه به حجم وسیع اطلاعات در وب، مدیریت آن با ابزارهای سنتی تقریبا غیر ممکن است و ابزارها و روش هایی نو برای مدیریت آن مورد نیاز است [۱].

^۱ cimmment

زنده متولد می‌شوند، رشد می‌کنند و می‌میرند پس دارای منحنی عمر یا چرخه‌ی عمر هستند.

در این مقاله نشان داده شده که با ترکیب بازدیدهای وبسایت و واکنش شبکه‌های اجتماعی می‌توان مقاله‌های خبری را دسته‌بندی کرد. همچنین می‌توان از واکنش شبکه‌های اجتماعی آنلاین برای پیش‌بینی الگوی بازدیدها استفاده کرد.

اصطلاح رده در رسانه به گروه‌ها یا اداره‌های خبری (سرویس‌های خبری) مانند رده‌ی ورزشی، رده‌ی سیاسی و غیره گفته می‌شود. معمولاً هر خبرگزاری یا وبسایت خبری اخبار خود را در رده‌هایی رده‌بندی می‌کند. در این کاربرد، رده‌بندی اخبار نوعی دسته‌بندی است و اخبار در کلاس‌های از پیش تعیین‌شده‌ای (رده‌ها) قرار داده می‌شوند. می‌توان با استفاده از روش‌های داده کاوی چرخه عمر رده‌های خبری را در شبکه‌های اجتماعی مورد تحلیل قرار داد. مطابق آمار سایت الکسا (۲۰۱۲) ^۲ فیس بوک و توئیتر یکی از ۱۰ سایت برتر مورد استفاده ایرانیان در فضای وب است [۳]. شبکه‌های اجتماعی مانند فیس بوک و توئیتر دارای ۵۰۰ میلیون کاربر در سراسر جهان هستند. آنها با هم در حال ایجاد یک جامعه از مصرف‌کنندگان، دوستان، خانواده، همکاران و آشنایان هستند که به بیش از ۳۰ میلیارد قطعه از محتوا (لینک‌های وب، اخبار، پست‌های وبلاگ، یادداشت‌ها، آلبوم عکس، و غیره) در هر ماه به اشتراک می‌گذارند [۴]. رسانه‌های اجتماعی، به علت افزایش آگاهی، به اشتراک‌گذاری اطلاعات، تشکیل نظرات، تبدیل به یک عامل نفوذی در رفتار مصرف‌کنندگان شده است [۵]. امروز بسیاری از وبسایت‌ها دکمه‌های فیس بوک و توئیتر و ... را در وبسایت جاسازی می‌کنند، که نه تنها اجازه می‌دهد تا مصرف‌کنندگان به تبادل اطلاعات در شبکه‌های اجتماعی در یک سرعت بسیار سریعتر بپردازند؛ بلکه شرکت‌ها را برای توجه بیشتر به واکنش طرفداران از طریق رسانه‌های اجتماعی متمایل می‌کند. حمایت طرفداران شبکه‌های اجتماعی از پیام‌های یک شرکت می‌تواند یک شاخص مهم از چگونگی اثربخشی استراتژی‌های بازاریابی اجتماعی یک شرکت باشد [۶].

حال با توجه به اهمیت شبکه‌های اجتماعی و تاثیر آن در تمام حوزه‌ها قصد داریم محتوای خبری را در شبکه‌های اجتماعی ردیابی کنیم در واقع بازتاب شبکه‌های اجتماعی مانند فیس بوک و توئیتر را روی موضوعات منتشر شده وبسایتها بررسی کنیم. با بررسی نتایج پژوهش‌ها می‌توان به تفاوت‌ها و شباهت‌های محتوای خبری در کاربردهای وبکاوی پی برد.

۱-۱- ویژگی‌های مقاله‌های خبری برخط

امروزه در یک صفحه‌ی وبسایت خبری علاوه بر متن خبری واحدهای بسیاری حاوی اطلاعات تکمیلی نظرخواهی‌ها، نظرها، تاریخ، ساعت، پیشنهادها، پیوندها و غیره وجود دارند [۷]. به طور کلی می‌توان گفت محتوای خبری به چند بخش قابل تقسیم است و معمولاً شامل مواردی چون عنوان، چکیده، نویسنده‌ها، تاریخ انتشار، پیوندهای مرتبط و غیره می‌شود [۸]. صفحه‌های خبری برخط دارای ساختاری قاعده‌مند متشکل از یک یا چند قالب ثابت با متن‌های پیوسته هستند [۹].

مقاله‌های خبری انواع مختلفی را شامل می‌شوند. در اخبار برخط مانند اخبار سنتی باید بتوانند به شش پرسش پایه‌ای «چه چیزی، چه کسی، چطور، چه

زمانی، کجا و به چه علت» ^۴ پاسخ دهند که به آن اجزای سازنده‌ی خبر می‌گویند [۱۰، ۱۱]. این ویژگی منحصر به شش پرسش نبوده و پژوهش‌گران پرسش‌های دیگری را نیز به آن اضافه کرده‌اند. برای مثال «برای چه کسی» ^۵ [۱۲]. و این که آیا منبع معتبری در خبر ذکر شده است و کاربران تا چه سطحی در تولید آن مشارکت داشته‌اند [۴]. برخی از ویژگی‌های مقاله‌های خبری برخط که در پژوهش‌های پیشین شناسایی شده‌اند در جدول (۱) آورده شده است.

جدول (۱): ویژگی‌های وبسایت‌های خبری شناسایی شده در

پژوهش‌های پیشین

ویژگی	منبع
ساختار قاعده‌مند	[۶، ۱۳]
ارزش‌های مکانی متفاوت	[۱۵، ۱۴]
محتوای پویا	[۱۶]
فضای نامحدود	[۱۷]
تاملی بودن	[۱۸، ۴] [۱۹]

۲- پیشینه

در جدول (۲) خلاصه‌ی شاخص‌های حیاتی مقاله‌های خبری بررسی شده در پژوهش‌های پیشین آورده شده است.

جدول (۲): شاخص‌های حیاتی مقاله‌های خبری بررسی شده در

پژوهش‌های پیشین

منبع	روش	شاخص	متغیر
کاس-تیلو و همکارانش (۱۱)	اثرگذاری اجتماعی	تعداد بازدیدها	کل بازدیدها منتهای ارجاع‌های داخلی
بی‌یو و همکارانش (۱۲)	اثرگذاری اجتماعی	پیامهای ارسال شده	بها اشتراک‌گذاری‌ها در دقیقه در فیسبوک
ونگ و همکارانش (۱۳)	تحلیل شبکه	اقتباس‌ها	تعداد مقاله‌های منتشر شده
			تعداد وبسایت‌های اقتباس‌کننده
ساکایس (۱۴)	اندازه‌گیری اهمیت و اثرگذاری مقاله خبری	تعداد نظرها	تعداد مقاله‌های منتشر شده
			تعداد مقاله‌های دارای نظر
			تعداد نظرها
			زمان نظرگذاری‌ها
زنگ زیانگ و همکارانش (۲۰)	اثرگذاری اجتماعی	بها اشتراک گذاشتن نظرات و تجربیات شخصی	با استفاده از مجموعه‌ای از کلمات کلیدی از پیش تعریف شده
هیما بندی و همکارش (۲۱)	پیش‌بینی توجه در رابطه با یک پست	تعداد نظرات دریافت شده توسط یک پست	ویژگی محتوا و ویژگی‌های سطح جامعه
جی وان و همکارش (۲۲)	اثرگذاری اجتماعی	پستهای توئیتر	محبوبیت قطعات محتوا در وب

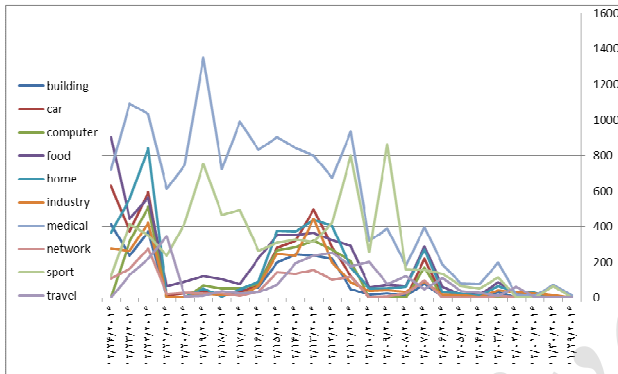
^۴ What, Who, How, When, Where and Why (Six Ws) For whom ^۵

^۲ News Categorization alexa ^۳

جدول (۴) : فراوانی بالاترین رده های خبری

رده	تعداد
سرگرمی و اجتماعی	۶۴۶۱۴
پزشکی	۱۴۸۲۲
ورزشی	۸۴۰۸

اما نکته قابل توجه زمان مرگ رده های خبرهای در تاریخ ۲۰۱۴/۱۱/۲۸ و شروع تولد مجدد آنها است. اما برای رده سرگرمی و اجتماعی طول عمر خبرها بیشتر است و در این بازه زمانی نمودار خط افق را قطع نمی کند در حالیکه برای سایر رده ها این اتفاق رخ می دهد. ولی در روزهای پایانی ماه نمودار رده های پر بازدید نیز به سمت افق نزدیک و مرگ آنها فرا می رسد. برای نشان دادن این مطلب رده پر قدرت را حذف تا وضعیت سایر رده ها مشخص شود. همان طور که در شکل (۵) مشخص است می توان گفت طول عمر رده های پر بازدید شامل سرگرمی و اجتماعی، پزشکی و ورزشی نسبت به سایرین بیشتر است. سپس رده های غذا، خانگی، ماشین، صنعت، کامپیوتر، ساختمان و سفر و رده شبکه به ترتیب کمترین طول عمر را داشته است.



شکل (۶) : نمودار فراوانی رده های خبری

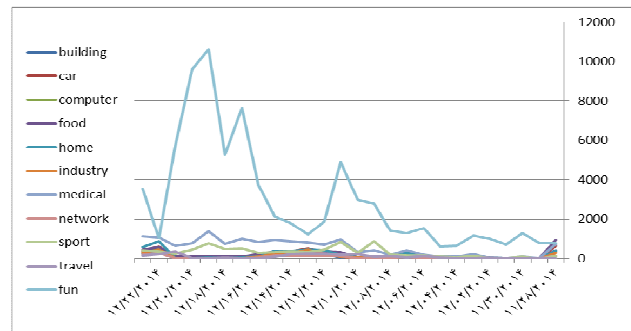
۴- یافته ها و ارزیابی داده

با استفاده از تجزیه و تحلیل کیفی و همچنین تجزیه و تحلیل کمی در داده ها از یک شبکه بزرگ خبر، برای مجموعه ای از داده های ثبت شده^۹ بیش از ۱۰۰۰۰۰۰ بازدیدکننده که پس از پاکسازی و حذف رکوردهای بدون مقدار و فیلتر سازی تمامی کدهای خبر کلیک شده از طرف شبکه های اجتماعی دارای ۱۲۱۳۶۶ رکورد به دست آمد. که هر خبر دارای ویژگی های متعددی از جمله تاریخ خبر، زمان کلیک خبر، کد خبر، عنوان خبر، رده خبر، ایام هفته، تعداد کلیک و... بودند که با استفاده از دسته بندهای درخت تصمیم از جمله J۴۸، Decision Tree و متغیر هدف تعداد کلیک با استفاده از نرم افزار ریپید ماینر خروجی حاصل شد که با توجه به این نوع الگوریتم های داده کاوی به کمک نرم افزار « ریپید ماینر»، « دقت»^{۱۰} اندازه گیری شد که دقت^{۱۱} هر کدام طبق جدول (۶) محاسبه، که دقت J۴۸ نسبت به سایر روشها بالاتر بود. بنابراین با دقت ۹۸.۸۸ درصد می توان تعداد کلیک هر رده خبر و نهایتاً طول عمر خبر را پیش بینی کرد.

پیک توجه منابع خبری	ریتیم چرخه اخبار به صورت روزانه	ردیابی الگوی رفتاری	لسکوک و همکارانش (۲۳)
زمان تاخیر	اشتراک گذاری محتوا از طریق انتشار اسناد	اثرگذاری اجتماعی	نیانا و همکارانش (۲۴)

۳- روش تحقیق

با توجه به اینکه سایت ایران نیدز^۱ دارای تعداد بازدید کننده روزانه حدوداً ۱۰۰۰۰۰ بازدید کننده می باشد و یک سایت خبری با ۱۹ رده خبری متفاوت که با توجه به افزایش داده های خبری را به ۱۱ عنوان شامل: سرگرمی و اجتماعی، پزشکی، ورزشی، تغذیه، اتومبیل، خانگی، صنعت، کامپیوتر، سفر، شبکه و ساختمان کاهش دادیم. وقایع ثبت شده^۷ از این سایت از تاریخ ۲۰۱۴/۱۱/۲۸ الی ۲۰۱۴/۱۲/۲۳ به مدت ۲۶ روز مورد تحلیل قرار گرفت. مجموع داده پس از پاکسازی و حذف داده های غیر منطقی و بدون مقدار به ۱۰۰۰۰۰۰ رکورد تبدیل شد. این داده ها شامل ویژگی هایی از قبیل تاریخ، زمان، کد خبر، رده خبر، آدرس ارجاع، روز هفته، هدف از فعالیت، وضعیت پرس و جو، کد وضعیت ویندوز، تعداد بایت ارسالی به سرور و تعداد بایت ارسالی به کاربر و ... می باشد. با توجه به اینکه تاثیر شبکه های اجتماعی روی چرخه عمر خبر مدنظر بود؛ داده هایی که آدرس ارجاع آنها شبکه های اجتماعی فیس بوک و توییتر بود مورد انتخاب قرار گرفت؛ که با احتساب همه شرایط تعداد داده مورد تحلیل به ۱۲۱۷۶۳ داده رسید. ابتدا با استفاده از نرم افزار ریپید ماینر فراداده^۸ موجود را بررسی کرده و به تفکیک، فراوانی هر رده خبر را نسبت به هر روز از بازه زمانی به دست آورد. شکل (۳) نمودار فراوانی رده ها را نشان می دهد.



شکل (۳) : نمودار فراوانی رده های خبری

همانطور که در شکل مشخص است رده سرگرمی و اجتماعی بیشترین فراوانی را به خود اختصاص داده است و نسبت به سایر رده ها بیشترین کلیک را از طرف طرفداران خود داشته است. و بعد از رده سرگرمی و اجتماعی رده های پزشکی و ورزشی بیشترین فراوانی را دارند و سپس سایر رده ها. در جدول (۴) فراوانی بالاترین رده های خبری را نشان می دهد.

^۹ log

^{۱۰} دقت: عبارت است از تعداد نمونه هایی که به درستی تشخیص داده می شوند، نسبت به کل

نمونه ها

^{۱۱} accuracy

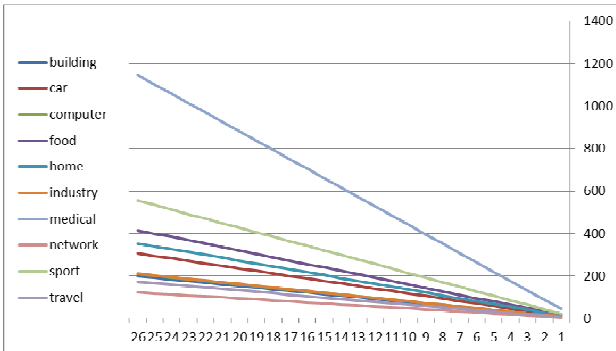
^۱ http://www.iranneeds.com

^۷ log

^۸ Meta date

شکل (۸): نمودار شیب خط رده ها

که می توان مجدد رده پر بازدید را حذف و وضعیت سایر رده ها را در شکل (۹) بررسی کرد.



شکل (۹): نمودار شیب خط رده ها به غیر از سرگرمی و اجتماعی

همانطور که مشخص است رده های پر بازدید شیب بیشتری دارند پس به طور متوسط با نرخ کمتری به خط افق که همان زمان مرگ خبر است نزدیک می شوند و می توان گفت طول عمر بیشتری دارند.

۵- نتیجه

هدف اصلی به کارگیری الگوریتم های داده کاوی در بررسی چرخه عمر خبر و بهره گیری بهتر از پایگاه داده و کشف دانش است. نکته قابل توجه این است که میزان موفقیت این الگوریتم ها، به عوامل زیادی بستگی دارد؛ عواملی مانند تعداد متغیر ها، بزرگی پایگاه داده، تعداد داده های پرت و مفقود، عدم دسترسی به داده های مناسب و درست در نتیجه داده کاوی موثر است؛ و دارای دقت و حساسیت و ویژگی متفاوتی هستند. پس جهت کسب نتیجه مناسب، در ابتدا باید از صحت و درستی اطلاعات مطمئن شویم و سپس از الگوریتم مناسب جهت داده کاوی استفاده نماییم. در این تحقیق روش های متعدد دسته بندی بر پایه درخت تصمیم جهت ارزیابی Accuracy مورد آزمایش قرار گرفته شد. همچنین با استفاده از رگرسیون خطی تابع تخمین انواع رده ها حاصل شد و نمودار رده ها که شیب خط، طول عمر خبر را ردیابی می کرد؛ به دست آمد. نتایج نشان داد که به ترتیب رده های سرگرمی، پزشکی، ورزشی، غذا، ماشین، خانگی، صنعت، سفر، ساختمان، کامپیوتر و شبکه دارای محبوبیت می باشند که محبوبیت خبر به رده خبری بستگی دارد؛ یعنی خبرهای رده سرگرمی و اجتماعی دارای نیمه عمر بالاتر و سپس اخبار پزشکی و ورزشی و رده خبری شبکه کمترین طول عمر را دارا بودند.

۶- مراجع

- [۱] R. Kosala, H. Blockeel. "Web mining research : A survey". In SIGKDD Explorations Newsletter of the ACM Special Interest Group on Knowledge Discovery and Data Mining, ۲(۱), ۲۰۰۰, ۱-۱۰

جدول (۶). درصد دقت انواع دسته بندی ها

J۴۸	Random Forest	ID۳	Decission Tree	Accuracy
۹۸.۸۸	۹۸.۷۷	۹۵.۸۷	۶۲.۱۰	

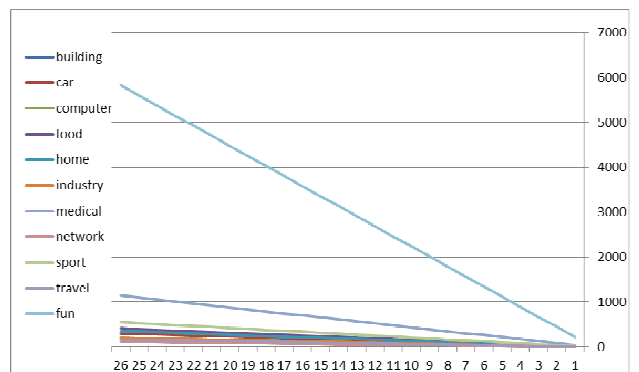
۴-۱- تابع تخمین برای تمام رده های خبری

بر اساس الگوریتم رگرسیون در محیط spss ضریب همبستگی بین تعداد خبر کلیک شده و سایر ویژگی ها بررسی وضریب همبستگی بین تعداد خبر تعداد روز قابل ملاحظه بود. تابع به صورت $y=Bx+constant$ استخراج شد که تعداد کلیک رده های خبری $y=$ تعداد روز $x=$ و B ضریب رگرسیون است که همان ضریب همبستگی می باشد و $constant$ ضریب ثابت می باشد. با توجه به اینکه مقدار ضریب اطمینان برابر ۹۵٪ است مقادیر sig با مقدر کمتر از ۰/۰۰۵ معنی دار بوده که این مقدار برای ضریب ثابت معنی دار نیست و مقدار ضریب ثابت صفر در نظر گرفته می شود؛ همچنین این مقدار برای سایر ویژگی ها نیز مقدار معناداری نبود که می توان گفت مهمترین ویژگی برای طول عمر خبر تعداد کلیک از طرف کاربران می باشد چون مقدار این ضریب صفر است؛ نهایتا نتیجه تابع برای ۱۱ رده خبری در جدول (۷) نشان داده شده است.

جدول (۷): توابع رگرسیون رده ها

ردیف	نام رده	تابع
۱	تابع رگرسیون رده شبکه	$Y=۴.۷۴۷ X$
۲	تابع رگرسیون رده سفر	$Y=۶.۶۱۶ X$
۳	تابع رگرسیون رده ساختمان	$Y=۷.۶۶۶ X$
۴	تابع رگرسیون رده کامپیوتر	$Y=۸.۰۴۱ X$
۵	تابع رگرسیون رده صنعت	$Y=۸.۱۳۳ X$
۶	تابع رگرسیون رده ماشین	$Y=۱۱.۷۸۱ X$
۷	تابع رگرسیون رده خانگی	$Y=۱۳.۶۰۶ X$
۸	تابع رگرسیون رده غذا	$Y=۱۵.۹۹۲ X$
۹	تابع رگرسیون رده ورزش	$Y=۲۱.۳۵۴ X$
۱۰	تابع رگرسیون رده پزشکی	$Y=۴۴.۰۱۶ X$
۱۱	تابع رگرسیون رده سرگرمی و اجتماعی	$Y=۲۳۳.۸۵۲ X$

حال طبق توابع به دست آمده می توان طول عمر رده ها را بر اساس شیب خط هر کدام از رده ها بررسی کرد. همانطور که در شکل (۸) مشخص است شیب خط به طور متوسط نشان دهنده تعداد کلیک بر اساس تاریخ می باشد. پس خطهای شیب دار نشان دهنده تعداد کلیک بیشتر و طول عمر بیشتر است.



- [۲] تحلیل محتوای صفحات و گروه های شبکه اجتماعی فیس بوک با نظریه حوزه عمومی هابرماس"، مطالعات و تحقیقات اجتماعی در ایران / دوره دوم، شماره ۱، بهار ۱۳۹۲ : ۱۱۰-۸۷.
- [۳] سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران، "جستارهایی درباره آرشيو وب"، ۱۳۹۱،
- [۴] <http://www.fsbassociates.com>.
- [۵] W. G. Mangold, and D. J. Faulds, "Social media: The new hybrid element of the promotion mix", (۲۰۰۹), Business horizons ۵۲(۴): ۳۵۷-۳۶۵.
- [۶] B. Yu, et al, "Toward Predicting Popularity of Social Marketing Messages", Social Computing, Springer Berlin Heidelberg. (۲۰۱۱). ۶۵۸۹: ۳۱۷-۳۲۴.
- [۷] ا. ریسی و انانی و م. فرزانه، "بررسی جامع تاثیر خدمات پر کاربرد وبسایت های خبری بر کیفیت آن ها"، مجله ی مدیریت فناوری اطلاعات، vol. دوره ی ۳، شماره ی ۷، ۱۳۹۰، ۵۴-۳۷.
- [۸] A. Spengler and P. Gallinari, "Learning to Extract Content from News Webpages", International Conference on Advanced Information Networking and Applications Workshops, WAINA '۰۹, (۲۰۰۹).
- [۹] J. Yan, et al, "News Web Text Extraction Based on the Maximum Subsequence Segmentation", Fifth IEEE International Conference on Computational and Information Sciences (ICIS), (۲۰۱۳).
- [۱۰] S. Mfenyana, et al, "Development of a Facebook Crawler for Opinion Trend Monitoring and Analysis Purposes: Case Study of Government Service Delivery in Dwesa", (۲۰۱۳), Development ۷۹(۱۷).
- [۱۱] W. Xuan., et al, "A Novel Approach To Automatically Extracting Main Content of Web News", International Conference on E-Business and Information System Security, EBISS '۰۹, (۲۰۰۹).
- [۱۲] J. Chen, et al, "An adaptive bottom up clustering approach for Web news extraction", ۱۸th IEEE Annual Wireless and Optical Communications Conference, WOCC, (۲۰۰۹).
- [۱۳] C. Castillo, et al, "Characterizing the life cycle of online news stories using social media reactions", Proceedings of the ۱۷th ACM conference on Computer supported cooperative work & social computing. Baltimore, Maryland, USA, (۲۰۱۴), ۲۱۱-۲۲۳.
- [۱۴] W. Youzhong, et al, "Propagation of online news: Dynamic patterns", IEEE International Conference on Intelligence and Security Informatics, ISI '۰۹, (۲۰۰۹).
- [۱۵] M. Tsagkias, et al, "Predicting the volume of comments on online news stories", Proceedings of the ۱۸th ACM conference on Information and knowledge management, Hong Kong, China, (۲۰۰۹), ACM: ۱۷۶۵-۱۷۶۸.
- [۱۶] Z. Xiang, and U. Gretzel, "Role of social media in online travel information search." (۲۰۱۱), Tourism management ۳۱(۲): ۱۷۹-۱۸۸.
- [۱۷] J. G. Lee, et al, "An approach to model and predict the popularity of online contents with explanatory factors", IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), (۲۰۱۰).
- [۱۸] J. Yang, and J. Leskovec, "Patterns of temporal variation in online media", Proceedings of the fourth ACM international conference on Web search and data mining, (۲۰۱۱).
- [۱۹] J. Leskovec, et al, "Meme-tracking and the dynamics of the news cycle", Proceedings of the ۱۵th ACM SIGKDD international conference on Knowledge discovery and data mining, (۲۰۰۹).
- [۲۰] Z. Xiang, and U. Gretzel, "Role of social media in online travel information search." (۲۰۱۱), Tourism management ۳۱(۲): ۱۷۹-۱۸۸.
- [۲۱] H. Lakkaraju, and J, "Ajmera Attention prediction on social media brand pages", Proceedings of the ۲۰th ACM international conference on Information and knowledge management, (۲۰۱۱).
- [۲۲] J. Yang, and J. Leskovec, "Patterns of temporal variation in online media", Proceedings of the fourth ACM international conference on Web search and data mining, (۲۰۱۱).
- [۲۳] J. Leskovec, et al, "Meme-tracking and the dynamics of the news cycle", Proceedings of the ۱۵th ACM SIGKDD international conference on Knowledge discovery and data mining, (۲۰۰۹).
- [۲۴] S. Mfenyana, et al, "Development of a Facebook Crawler for Opinion Trend Monitoring and Analysis Purposes: Case Study of Government Service Delivery in Dwesa", (۲۰۱۳), Development ۷۹(۱۷).