

# ایجاد سایتهاي وب تطبیق‌پذیر با استفاده از اتوماتاهای یادگیر

علی برادران هاشمی<sup>۱</sup> محمد رضا میبدی<sup>۲</sup>

## چکیده

تطبیق‌پذیر کردن یک سایت وب موجب افزایش رضایت کاربران و افزایش استفاده از آن سایت وب می‌شود. در این مقاله یک روش جدید مبتنی بر اتوماتاهای یادگیر برای ایجاد یک سایت وب تطبیق‌پذیر با کمک داده‌های استفاده از وب پیشنهاد شده است. روش پیشنهادی با بهره‌گیری از رفتار کاربران در گذشته سعی در تغییر ساختار سایت وب دارد. بصورتی که تغییرات انجام شده منجر به کوتاه‌تر شدن مسیر کاربران به صفحات هدف شود. با کوتاه‌تر شدن مسیر کاربر در سایت وب و کاهش مدت زمان حضور هر کاربر در سایت وب بمنظور دستیابی به اطلاعات مورد نیاز خود، علاوه بر افزایش رضایت کاربران، ظرفیت پذیرش کاربران سایت نیز افزایش می‌یابد.

## کلمات کلیدی

داده‌کاوی استفاده از وب، اتوماتای یادگیر، سایتهاي تطبیق‌پذیر.

## Adaptive Web Sites Using Learning Automata

Ali B. Hashemi<sup>۱</sup> M. R. Meybodi<sup>۲</sup>

{a\_hashemi<sup>۱</sup>, mmeybodi<sup>۲</sup>}@aut.ac.ir

Soft Computing Lab, Computer Engineering & Information Technology Department,  
Amirkabir University of Technology, Tehran, Iran

## Abstract

Adaptive web sites have two advantages, increasing user satisfaction and higher utilization of the web site because of decreasing users dwell time. This paper proposes a novel learning automata based algorithm to make a web site adaptive by applying web usage data. The proposed algorithm changes web site structure according to the past user behavior. The goal is making access to the web site easier by reducing the length of the path traversed by the user. Computer simulations have been used to study performance of the algorithm. The results of simulations have been compared with some of existing methods reported in the literature.

## Keywords

Web usage mining, Learning automata, Adaptive web sites.

## ۱ مقدمه

با گسترش وب و کاربرد آن در دنیای تجارت، بسیاری از خدمات ارائه شده توسط شرکتها و موسسات از طریق وب انجام می‌پذیرد. بهمین دلیل طراحی سایتهاي وب به اندازه‌ای مهم می‌باشد که می‌تواند در ادامه حیات یک تجارت موثر باشد. سه پارامتر زیبایی، محتوا و ساختار یک سایت وب موفقیت یک سایت وب مشخص می‌کنند. با افزایش تعداد صفحات یک سایت وب نوع محتوا و قالب طراحی سایت تغییری نخواهد کرد

<sup>۱</sup> آزمایشگاه محاسبات نرم، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران، ایران, a\_hashemi@aut.ac.ir

<sup>۲</sup> استاد و عضو هیئت علمی، آزمایشگاه محاسبات نرم، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران، ایران,

mmeybodi@aut.ac.ir

اما مسلما ساختار سایت دچار تغییر خواهد شد. مساله اصلی برای حفظ کیفیت سایت و راضی نگهداشتن کاربران سایت، نحوه تغییر سایت و چگونگی قرار دادن صفحات جدید است [۱]. به این دلیل که اگر افزایش تعداد صفحات سایت بصورت مناسبی انجام نشود، کاربران برای پیدا کردن صفحات مورد نظرشان مدت زمان بیشتری را صرف خواهند کرد که موجب کاهش رضایت آنها می‌شود.

دو رویکرد برای تعیین ساختار مناسب یک سایت وجود دارد. رویکرد نخست، استفاده از راهکارهایی مانند بهره‌گیری از نظرات افراد خبره یا نظرسنجی از کاربران سایت، برای تغییر ساختار سایت وب است. در رویکرد دوم، با استفاده از روش‌های داده‌کاوی در داده‌های وب شامل ساختار سایت وب و داده‌های استفاده کاربران از سایت وب (فایلهای ثبت واقعی<sup>۱</sup> سرویس‌دهنده وب) ساختار سایت وب بر اساس رفتار کاربران و با تشخیص نیازهای اطلاعاتی آنها تغییر می‌دهند. تغییر محتوای و ساختار یک سایت بر اساس علایق هر کاربر یا دسته‌ای از کاربران، یا ایجاد سایتهای تطبیقی، یکی از مسائل مهم تجارت الکترونیک بشمار می‌رود. بطوریکه ارائه دقیق اطلاعات یا محصولات مورد نیاز کاربران می‌تواند به طرز چشمگیری کارایی سایت وب را افزایش دهد. Perkowitz و Etzioni در [۲]، [۳]، [۴] و [۵] با خوشبندی داده‌های استفاده کاربران از یک سایت وب فهرست پویایی از صفحات سایت ایجاد می‌کنند که با تغییر رفتار کاربران سایت و علایق آنها خود را تطبیق می‌دهد. کارایی این روش با پیاده‌سازی واقعی آن در یک سایت وب و با اندازه‌گیری کاربرانی که از فهرست پویای تولیدشده استفاده می‌کرند نشان داده شده است. Lee و Shiu [۶] با معرفی معیاری بنام کارایی سایت وب الگوریتمی برای تغییر ساختار یک سایت وب با هدف افزایش کارایی سایت وب و سادگی دسترسی کاربران ارائه کرده‌اند. این روش با مشاهده رفتار کاربران یک سایت وب، بین صفحاتی از سایت که تشخیص داده می‌شود از نظر کاربران مهم هستند لینکی قرار می‌دهد یا پیشنهاداتی برای تغییر طراحی صفحات با کارایی پایین ارائه می‌کند. Theodoulidis و Mikroyannidis در [۷] چارچوبی برای ایجاد یک سایت وب پویا که با استفاده از داده‌های استفاده از وب و هستان‌شناسی ساختار خود را با رفتار کاربران تطبیق می‌دهد ارائه کرده‌اند. Warner و همکاران در [۸] نتایج پیاده‌سازی یک سایت پویا با استفاده از روش کلونی مورچه‌ها را گزارش کرده‌اند. Shokry و همکاران در [۹] روش جدید با استفاده از خوشبندی فازی برای ایجاد صفحات فهرست پویا ارائه کرده‌اند که نتایج شبیه‌سازی این روش کارایی پیشتری نسبت به روش [۲] نشان می‌دهد. ایده روش Srikant و Yang [۱۰] این است که نقطه‌ای در گراف حرکت کاربر که در آن کاربر از یک مسیر انتخاب شده بازگشت می‌کند، مکان صحیح قرار گرفتن صفحه مورد نظر وی می‌باشد. با تشخیص چنین محلی، پیشنهاد قرار دادن یک لینک به صفحه مورد نظر به مدیر سایت وب داده می‌شود. این روش بر روی یک سایت وب کوچک اجرا شده و نتایج بدست آمده نشان می‌دهد که حتی در یک سایت وب کوچک نیز چنین نقاط وجود دارد. در روش Weigang و همکاران [۱۱] ساختار یک سایت وب با استفاده از کلونی مورچه‌ها و با هدف افزایش تعداد کاربرانی که در یک بازه زمانی معین به صفحه هدف خود رسیده‌اند، تغییر داده می‌شود. Lin در [۱۲] با هدف افزایش استفاده کاربران از همه لینکها، تطبیقی کردن یک سایت وب را بصورت یک مساله بهینه‌سازی مدل کرده است.

در این مقاله یک روش جدید برای تغییر ساختار سایتهای وب با استفاده از اتماتای یادگیر پیشنهاد می‌شود از اتماتای یادگیر در کاربردهایی مانند شناسایی الگو [۱۳]، تشخیص منابع در شبکه‌های سلولی [۱۴] [۱۵] [۱۶] [۱۷] جستجو در وب [۱۸] و تعیین شباهت صفحات وب [۱۹] استفاده شده است. در الگوریتم‌های پیشنهادی از یک اتماتای یادگیر توزیع شده متناسب با ساختار اولیه سایت استفاده می‌شود. این اتماتای یادگیر با استفاده از داده‌های استفاده از سایت وب اقدام‌های خود را تنظیم می‌کند. بصورتیکه پس از مشاهده رفتار تعدادی از کاربران، می‌تواند با استفاده از وضعیت اتماتای یادگیر ساختار سایت را تغییر داد. نتایج شبیه‌سازی‌های انجام شده نشان می‌دهد که در ساختار جدید بدست آمده با استفاده از هر دو الگوریتم پیشنهادی، کاربران سایت با مشاهده تعداد صفحات کمتری به هدف خود می‌رسند که در نتیجه میزان مطلوبیت آنها از سایت وب افزایش خواهد یافت.

در ادامه در بخش دوم اتماتای یادگیر معرفی می‌شود. سپس در بخش سوم الگوریتم پیشنهادی برای تغییر خودکار ساختار سایت با استفاده از اتماتای یادگیر ارائه می‌شود. در بخش چهارم پس از معرفی مدل شبیه‌سازی، نتایج شبیه‌سازی الگوریتم‌های پیشنهادی نشان داده شده و در بخش ششم نتیجه‌گیری این پژوهش ارائه می‌گردد.

اتوماتای یادگیر یک مدل انتزاعی است که بطور تصادفی یک اقدام از مجموعه متناهی اقدام‌های خود را انتخاب کرده و بر محیط اعمال می‌کند. محیط اقدام انتخاب شده توسط اتوماتای یادگیر ارزیابی کرده و نتیجه ارزیابی خود را توسط یک سیگنال تقویتی به اتوماتای یادگیر اطلاع می‌دهد. سپس اتوماتای یادگیر با اطلاع از اقدام انتخاب شده و سیگنال تقویتی، وضعیت داخلی خود را بروز کرده و اقدام بعدی خود را انتخاب می‌کند. شکل ۱ نحوه ارتباط بین اتوماتای یادگیر و محیط را نشان می‌دهد.



محیط را می‌توان توسط سه‌تایی  $E = \{\alpha, \beta, c\}$  نشان داد که در آن  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  مجموعه ورودیها،  $\beta = \{\beta_1, \beta_2, \dots, \beta_r\}$  مجموعه خروجیها و  $c = \{c_1, c_2, \dots, c_r\}$  مجموعه احتمالات جریمه می‌باشد. هرگاه  $\beta$  مجموعه دو عضوی باشد، محیط از نوع P می‌باشد. در چنین محیطی  $\beta_1 = 1$  و  $\beta_2 = 0$  به عنوان پاداش در نظر گرفته می‌شود. در محیط از نوع Q، مجموعه  $\beta$  دارای تعداد متناهی عضو می‌باشد و در محیط از نوع S، تعداد اعضاء مجموعه  $\beta$  نامتناهی است.  $c_i$  نشان دهنده احتمال نامطلوب بودن سیگنال تقویتی محیط در پاسخ به اقدام  $\alpha_i$  می‌باشد. در یک محیط ایستا<sup>۳</sup> مقادیر  $c_i$  ها ثابت هستند، حال آنکه در یک محیط غیر ایستا<sup>۴</sup> این مقادیر در طی زمان تغییر می‌کنند. بر اساس اینکه تابع بروز رسانی وضعیت اتوماتای یادگیر (که با اطلاع از اقدام انتخاب شده و سیگنال تقویتی  $\beta$ ، وضعیت بعدی اتوماتای یادگیر را محاسبه می‌کند) ثابت یا متغیر باشد، اتوماتای یادگیر به دو دسته اتوماتای یادگیر با ساختار ثابت و اتوماتای یادگیر با ساختار متغیر تقسیم می‌گردد. در این مقاله از اتوماتای یادگیر با ساختار متغیر استفاده شده است که در ادامه معرفی می‌شود.

اتوماتای یادگیر با ساختار متغیر توسط چهارتایی  $\{\alpha, \beta, p, T\}$  نشان داده می‌شود که در آن  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  مجموعه اقدام‌های اتوماتای یادگیر،  $\beta = \{\beta_1, \beta_2, \dots, \beta_r\}$  مجموعه ورودیها اتوماتای یادگیر،  $p = \{p_1, p_2, \dots, p_r\}$  بردار احتمال انتخاب هر یک از اقدام‌ها و  $T$  الگوریتم یادگیری اتوماتای یادگیر می‌باشد. الگوریتم‌های یادگیری متنوعی برای اتوماتای یادگیر ارائه شده است که در ادامه یک الگوریتم یادگیری خطی برای اتوماتای یادگیر بیان می‌گردد. فرض کنید اتوماتای یادگیر در مرحله n ام اقدام  $\alpha_i$  خود را انتخاب نموده و محیط ارزیابی خود را توسط سیگنال تقویتی  $\beta(n)$  به اتوماتای یادگیر اعلام کند. با استفاده از الگوریتم یادگیری خطی، اتوماتای یادگیر بردار احتمال انتخاب اقدام‌های خود را مطابق رابطه (۱) تنظیم می‌کند.

$$\begin{aligned} p_i(n+1) &= p_i(n) + a.(1 - \beta(n)).(1 - p_i(n)) - b.\beta(n).p_i(n) \\ p_j(n+1) &= p_j(n) + a.(1 - \beta(n)).p_j(n) + \frac{b.\beta(n)}{r-1} - b.\beta(n).p_j(n) \quad \text{if } j \neq i \end{aligned} \quad (1)$$

اتوماتای یادگیری که در بالا معرفی شد، دارای تعداد اقدام‌های ثابتی می‌باشد. در بعضی از کاربردها به اتوماتای یادگیر با تعداد اقدام متغیر<sup>۴</sup> نیاز می‌باشد [۲۰]. یک اتوماتای یادگیر با تعداد اقدام متغیر، در لحظه n، اقدام خود را از یک زیر مجموعه غیر تهی از اقدام‌ها بنام مجموعه اقدام‌های فعال  $V(n)$  انتخاب می‌کند. انتخاب مجموعه اقدام‌های فعال اتوماتای یادگیر  $V(n)$  توسط یک عامل خارجی و بصورت تصادفی انجام می‌شود. نحوه فعالیت این اتوماتای یادگیر بصورت زیر است.

اتوماتای یادگیر برای انتخاب یک اقدام در زمان n ابتدا مجموع احتمال اقدام‌های فعال خود  $K(n)$  را محاسبه و بردار  $\hat{p}(n)$  را مطابق رابطه (۲) ایجاد می‌کند. آنگاه اتوماتای یادگیر یک اقدام از مجموعه اقدام‌های فعال خود را بصورت تصادفی و بر اساس بردار احتمال  $\hat{p}(n)$  انتخاب کرده و بر محیط اعمال می‌کند. در یک اتوماتای یادگیر با الگوریتم یادگیری خطی، اگر اقدام انتخاب شده  $\alpha_i$  باشد، اتوماتای یادگیر پس از دریافت

پاسخ محیط، بردار احتمال  $(\hat{p}(n))$  اقدامهای خود در صورت دریافت پاسخ مطلوب بر اساس رابطه (۳) و در صورت دریافت پاسخ نامطلوب طبق رابطه (۴) بروز می‌کند. سپس اتوماتای یادگیر بردار احتمال اقدامهای خود  $p(n)$  را با استفاده از بردار  $\hat{p}(n+1)$  و طبق رابطه (۵) بروز می‌کند.

$$K(n) = \sum_{\alpha_i \in V(n)} p_i(n) \quad (2)$$

$$\hat{p}_i(n) = prob[\alpha(n) = \alpha_i | \alpha_i \in V(n)] = \frac{p_i(n)}{K(n)} \quad (2)$$

$V(n)$  is the set of enabled actions

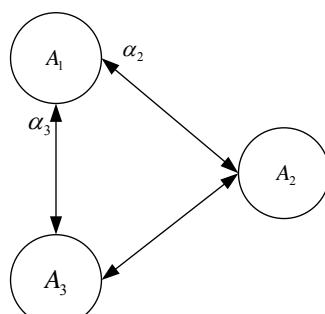
$$\begin{aligned} \hat{p}_i(n+1) &= \hat{p}_i(n) + a.(1 - \hat{p}_i(n)) \\ \hat{p}_j(n+1) &= \hat{p}_j(n) - a.\hat{p}_i(n) \quad \forall j \neq i \end{aligned} \quad (3)$$

$$\begin{aligned} \hat{p}_i(n+1) &= (1-b).\hat{p}_i(n) \\ \hat{p}_j(n+1) &= \frac{b}{r-1} + (1-b)\hat{p}_j(n) \quad \forall j \neq i \end{aligned} \quad (4)$$

$$\begin{aligned} p_i(n+1) &= \hat{p}_i(n+1).K(n) \quad \text{for all } i, \alpha_i \in V(n) \\ p_j(n+1) &= p_j(n) \quad \text{for all } j, \alpha_j \notin V(n) \end{aligned} \quad (5)$$

## ۲.۱ اتوماتای یادگیر توزیع شده

اتوماتای یادگیر توزیع شده [۲۱] شبکه‌ای از چند اتوماتای یادگیر است که برای حل یک مساله مشخص با یکدیگر همکاری می‌کنند. یک اتوماتای یادگیر توزیع شده را می‌توان با یک گراف جهت دار مدل کرد. بصورتی که مجموعه گره‌های آنرا مجموعه‌ای از اتوماتای یادگیر و یالهای خروجی هر گره مجموعه اقدامهای متناظر با اتوماتای یادگیر متناظر با آن گره است. هنگامی که اتوماتا یکی از اقدامهای خود را انتخاب می‌کند، اتوماتایی که در دیگر انتهای یال متناظر با آن اقدام قرار دارد، فعال می‌شود. بعنوان مثال در شکل ۲ هر اتوماتا ۲ اقدام دارد. اگر اتوماتای  $A_1$  اقدام  $\alpha_3$  خود را انتخاب کند، آنگاه اتوماتای  $A_3$  فعال خواهد شد. در گام بعد، اتوماتای  $A_3$  یکی از اقدامهای خود را انتخاب می‌کند که منجر به فعال شدن یکی از اتوماتاهای یادگیر متصل به  $A_3$  می‌شود. در هر لحظه فقط یک اتوماتای یادگیر در اتوماتای یادگیر توزیع شده فعال می‌باشد. بصورت رسمی، یک اتوماتای یادگیر توزیع شده با  $n$  اتوماتای یادگیر توسط یک گراف  $(A, E)$  تعریف می‌شود که  $A = \{A_1, A_2, \dots, A_n\}$  مجموعه اتوماتا و  $E \subset A \times A$  مجموعه لبه‌های گراف است بطوریکه لبه  $(j, i)$  متناظر با اقدام  $a_j$  از اتوماتای  $A_i$  است. اگر بردار احتمال اقدامهای اتوماتای یادگیر  $A_j$  با  $p^j$  نشان داده شود، آنگاه  $p_m^j$  احتمال انتخاب اقدام  $\alpha_m$  از اتوماتای یادگیر  $A_j$  را نشان می‌دهد که احتمال انتخاب لبه خروجی  $(j, m)$  از میان لبه‌های خروجی گره  $j$  می‌باشد.



شکل ۲. اتوماتای یادگیر توزیع شده

### ۳ الگوریتم پیشنهادی برای تعیین ساختار یک سایت وب

الگوریتم پیشنهادی با کاوش در داده‌های استفاده از یک سایت وب بدنال کاهش مسیر دسترسی کاربران به صفحات مورد نظر دارد. از آنجایی که ساختار بهینه برای یک سایت وب ناشناخته است، در الگوریتم پیشنهادی از یکی از روش‌های یادگیری تقویتی بنام اتوماتای یادگیر توسعه شده برای یادگیری ساختار بهینه یک سایت وب مطابق با رفتار کاربران استفاده می‌شود. نتیجه اجرای الگوریتم پیشنهادی ساختار جدیدی برای سایت وب مورد بررسی است که می‌توان بصورت خودکار یا دستی در سایت وب اعمال گردد. ساختار الگوریتم پیشنهادی بصورت زیر می‌باشد.

در یک سایت وب با  $n$  صفحه، از یک اتوماتای یادگیر توسعه شده [۲۱] که دارای  $n$ /اتوماتای یادگیر با تعداد اقدام‌های متغیر [۲۰] استفاده می‌شود. هر یک از اتوماتاهای استفاده شده  $n-1$  اقدام دارد که در هر لحظه تنها یک زیرمجموعه از این اقدامها فعال و قابل استفاده است. اگر در صفحه  $i$  لینکی به صفحه  $j$  وجود داشته باشد، آنگاه اقدام  $\alpha_j^i$  در اتوماتای متناظر با این صفحه  $LA_i$  فعال می‌باشد. به این ترتیب اتوماتای یادگیر توسعه شده بر اساس ساختار موجود در سایت مورد نظر ایجاد می‌شود. در الگوریتم پیشنهادی فرض شده است که حرکت کاربران در سایت وب تنها با استفاده از لینکهای موجود در هر صفحه انجام می‌شود. با در اختیار داشتن داده‌های استفاده از سایت وب برای هر یک از کاربران، شامل صفحاتی که یک کاربر از زمان ورود به سایت تا هنگام خروج مشاهده کرده است، الگوریتم پیشنهادی بصورت زیر عمل می‌کند.

#### Procedure Web\_Site\_Restructuring

**begin**

```

for all sessions do
    for all traversed links in this session such as  $(i,j)$  do
        Reward corresponding action to link  $j$  ( $\alpha_j^i$ ) in  $LA_i$  of page  $i$  ( $LA^i$ );
    end
end
// adapting  $LA$  to user behavior
for all actions of DLA do
    if the action probability <  $\tau_1$  then
        disable the action ;
    end
end
// restructuring web site
for all actions of DLA do
    if the action  $\alpha_j^i$  is enabled then
        put a link to page  $j$  in page  $i$ ;
    else
        remove link to page  $j$  from page  $i$ ;
    end
end

```

**end**

شکل ۳. شبکه الگوریتم پیشنهادی

با هر حرکت کاربر از صفحه  $i$  به صفحه  $j$ ، اتوماتای یادگیر متناظر با صفحه  $i$  ( $LA_i$ ) به اقدام متناظر با صفحه  $j$  ( $\alpha_j^i$ ) خود پاداش می‌دهد. این عمل برای تمام حرکات و کلیه کاربران سایت انجام می‌شود. پس از پایان یافتن این مرحله و بررسی کلیه لینکهای پیمایش شده، قسمت دوم الگوریتم برای انجام تغییرات در ساختار سایت آغاز می‌شود. در این قسمت با استفاده از وضعیت اتوماتای یادگیر (اقدام‌های فعال و بردار احتمال آنها) ساختار سایت بصورت زیر تغییر می‌کند. کلیه اقدام‌های اتوماتا که احتمال انتخاب آنها از یک مقدار آستانه‌ای  $\tau_1$  کمتر می‌باشد غیرفعال می‌شوند. در این حالت اگر اقدام زار اتوماتای  $i$  غیر فعال شده باشد، اتصال موجود از صفحه  $i$  به صفحه  $j$  برداشته می‌شود. با این کار انتخابهای یک کاربر در هر صفحه برای انتخاب صفحه بعد به صفحات مناسب‌تر محدود می‌گردد. در نتیجه از انتخاب صفحاتی که منجر به افزایش بیهوده حرکات کاربران در سایت می‌شود جلوگیری بعمل می‌آید. پس از اینکه این کار برای همه صفحات سایت وب انجام شد، ساختار سایت وب به این صورت تغییر می‌کند که متناظر با هر اقدامی فعال مانند  $\alpha_j^i$ ، لینکی از صفحه  $i$  به صفحه  $j$  قرار داده می‌شود. به این ترتیب با اجرای الگوریتم

پیشنهادی، لینکهایی که زیاد مورد توجه کاربران نمی‌باشند، با هدف کاهش انتخابهای سهولت انتخاب لینک بعدی و انتخاب سریعتر کاربران حذف می‌شود. شبیه کد این الگوریتم در شکل ۳ آمده است.

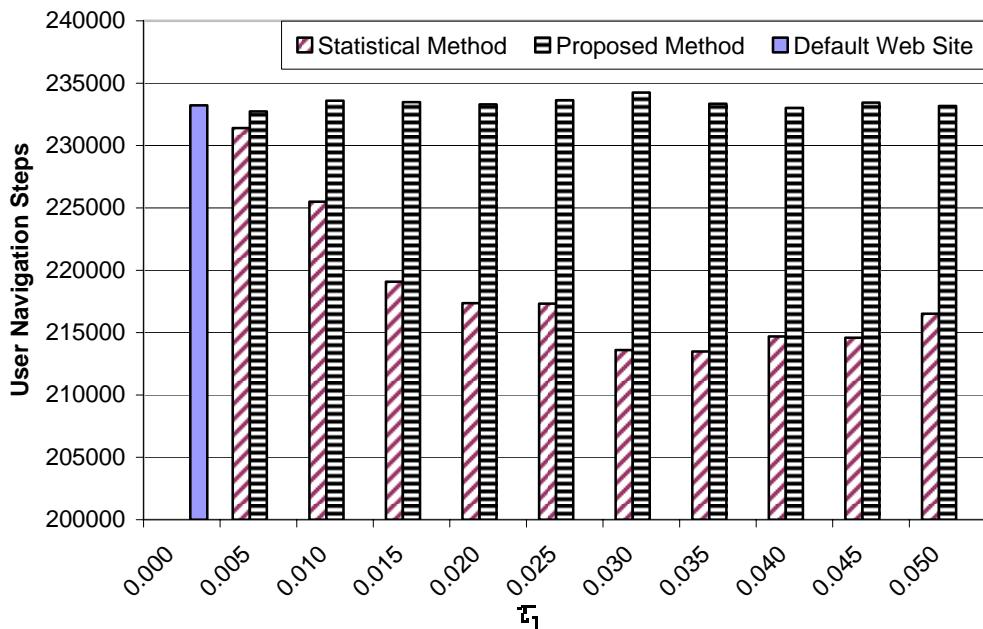
## ۴ نتایج شبیه‌سازی

در اکثر روش‌های گزارش شده در زمینه کاوش داده‌های استفاده از وب، از داده‌های یک سایت وب واقعی که برای محققین در دسترس بوده است (مانند سایت وب دانشگاه یا دانشکده)، استفاده شده است که بجز تعداد اندکی که از داده‌های [۲۳، ۲۲] استفاده کردند، این داده‌های در دسترس نمی‌باشند. انجام آزمایشات با چنین داده‌هایی علاوه بر تحمیل خطای ناشی از عملیات پیش‌پردازش، امکان تحت تاثیر قرار گرفتن<sup>۵</sup> الگوریتم‌های مورد استفاده بر اساس رفتار کاربران یک یا چند سایت خاص نیز ممکن می‌سازد. به دلایل فوق الذکر در روش انجام این پژوهش، با استفاده از مدل یک سایت وب و کاربران آن که توسط Liu و همکاران ارائه شده است [۲۴]، داده‌های استفاده از یک سایت وب فرضی تولید شده و در ارزیابی الگوریتم پیشنهادی استفاده می‌شود. مقادیر اختیار شده برای متغیرهای این مدل در جدول ۱ ذکر شده است.

جدول ۱ پارامترهای استفاده شده در مدل شبیه‌سازی

	آستانه ایجاد اتصال
۰/۷	
۵۰۰۰	تعداد کاربران
۲۶	تعداد اسناد
۴	تعداد موضوع‌ها
۰/۲	مقدار ثابت سند اولیه (صفحه اولیه سایت) در موضوعات مختلف $T_c$
-	ضریب ثابت کاهش اشتیاق کاربر $\Delta M_t^c$
-	ضریب متغیر کاهش اشتیاق کاربر $\Delta M_t^v$
۱	پارامتر توزیع Power-law توزیع احتمال علاقه کاربران $\alpha_u$
۱/۲	ضریب پاداش دریافی از مشاهده یک سند $\phi$
۰/۵	ضریب جذب اطلاعات از یک سند توسط یک کاربر $\lambda$
۵/۹۷	میانگین توزیع نرمال $\Delta M_t^v$
۰/۲۵	واریانس توزیع نرمال $\sigma_m$
-	میانگین توزیع نرمال برای مقدار افزایش یک گره برای یک موضوع خاص $\mu$
۳	پارامتر توزیع Power-law توزیع احتمال وزنهای مطالب برای هر سند $\alpha_p$
۰/۲۵	واریانس توزیع نرمال برای مقدار افزایش یک گره برای یک موضوع خاص $\sigma_t$
۱	ضریب کاهش علاقه کاربر $\theta$
۰/۲	حداقل اشتیاق کاربر برای ادامه جستجو

در مدل معرفی شده در [۲۴] برای ساختار یک سایت وب صفحاتی که فاصله بردار موضوعی آنها از آستانه ایجاد اتصال کمتر است به یکدیگر لینک خواهند داشت. برای نشان دادن کارایی الگوریتم پیشنهادی، مقدار آستانه ایجاد اتصال را از مقدار معرفی شده در [۲۴] (۰/۷) کمتر در نظر گرفته می‌شود (۰/۴). به این ترتیب تعداد لینکهای داخلی سایت وب تولید شده از حد مطلوب کمتر خواهد بود. با اجرای الگوریتم پیشنهادی بر روی داده‌های استفاده از این سایت وب، روش پیشنهادی می‌تواند لینکهای حذف شده (نسبت به مقدار ۰/۷) را تشخیص دهد. همانطور که در جدول زیر مشاهده می‌شود، با اجرای الگوریتم پیشنهادی و برای مقادیر مختلف  $\{\alpha_1 = 0.015, 0.020, 0.025, 0.030, 0.035, 0.040, 0.045, 0.05\}$  تعداد لینکهای سایت از ۸۵٪ تا ۵۳٪ کاهش می‌یابد. این کاهش موجب می‌شود کاربران با حرکات کمتری در سایت بتوانند صفحات مورد نظر خود را پیدا کرده و درخواستهای آنها در سایت وب تا ۸٪ کاهش یابد.



شکل ۴. تعداد حرکات کاربران در سایت وب با مقادیر مختلف  $\tau_1$  قبل از تغییر سایت وب

## ۵ نتیجه‌گیری

عموماً مدل ذهنی طراحان یک سایت وب با کاربران آن سایت متفاوت است. بهمین دلیل برای نگهداری یک سایت وب کاره، لازم است که نظرات کاربران برای طراحی ساختار سایت وب مورد توجه قرار گیرد. در غیر اینصورت کاربران برای دسترسی به صفحات مورد علاقه و اطلاعات مورد نیاز خود در آن سایت مدت زمان زیادی صرف می‌کنند که موجب نارضایتی آنها خواهد شد. یک رویکرد دریافت صریح نظرات کاربران با استفاده از پرسشنامه‌هایی است که در سایت قرار داده می‌شود که بعلت تحمیل یک وظیفه غیر ضروری از دیدگاه کاربر مناسب نمی‌باشد. رویکرد دیگر دریافت ضمنی نظرات کاربران با بررسی رفتار آنها در سایت وب است. در این مقاله یک الگوریتم جدید مبتنی بر اتماتاهای یادگیر برای دریافت ضمنی نظرات کاربران پیشنهاد شده است. الگوریتم پیشنهادی فقط با استفاده از داده‌های ثبت وقایع در یک سرویس‌دهنده وب می‌تواند ساختار جدیدی برای سایت وب مورد نظر پیشنهاد دهد، به صورتیکه این ساختار جدید موجب کاهش گامهای دسترسی کاربران به اطلاعات مورد نیاز شده و افزایش رضایت کاربران را بدنبل خواهد داشت. نتایج شبیه‌سازی الگوریتم پیشنهادی نشان داده است که با اجرای این الگوریتم می‌توان تعداد درخواستهای کاربران از یک سایت وب را تا ۸/۲٪ کاهش داد. این کاهش نشان دهنده افزایش سرعت دسترسی کاربران به اطلاعات و صفحات مورد نیاز می‌باشد که علاوه بر افزایش رضایت کاربران، موجب کاهش استفاده از منابع سرویس‌دهنده وب نیز می‌شود.

## ۶ مراجع

- [۱] T. Nakayama, H. Kato, and Y. Yamane, "Discovering the gap between Web site designers' expectations and users' behavior," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 33, no. 1-6, pp. 811-822, 2000.
- [۲] M. Perkowitz and O. Etzioni, "Adaptive Web sites," *Communications of the ACM*, vol. 43, no. 8, pp. 152-158, 2000.
- [۳] M. Perkowitz and O. Etzioni, "Adaptive Web Sites: an AI Challenge," in *Fifteenth International Joint Conference on Artificial Intelligence (IJCAI 97)*, 1997, pp. 16-23.
- [۴] M. Perkowitz and O. Etzioni, "Adaptive Web Sites: Conceptual Cluster Mining" in *17th International Joint Conference on Artificial Intelligence(IJCAI '99)*, Stockholm, Sweden, 1999.
- [۵] M. Perkowitz and O. Etzioni, "Towards adaptive Web sites: conceptual framework and case study," *Artificial Intelligence*, vol. 118, no. 1-2, pp. 245-275, 2000.

- [۱] J.-H. Lee and W.-K. Shiu, "An adaptive website system to improve efficiency with web mining techniques," *Advanced Engineering Informatics*, vol. 18, pp. 129–142, 2004.
- [۲] A. Mikroyannidis and B. Theodoulidis, "A Theoretical Framework and an Implementation Architecture for Self Adaptive Web Sites," in *International Conference on Web Intelligence*, 2004, pp.558-561.
- [۳] D. Warner ,S. D. Durbin, J. N. Richter, and Z. Gedeon, "Adaptive web sites: user studies and simulation," in *15th international conference on World Wide Web*, Edinburgh, Scotland, 2006, pp.975-976.
- [۴] R. A. Shokry, A. A. Saad, N. M. El-Makkey, and M. A. Ismail, "Using New Soft Clustering Technique in Adaptive Web Site," in *IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology*, 2006, pp.281-286.
- [۵] R. Srikant and Y. Yang, "Mining web logs to improve website organization," in *10th international conference on World Wide Web*, Hong Kong, Hong Kong, 2001 pp.430-437.
- [۶] L. Weigang, M. V. P. Dib, W. M. Teles, V. M. d. Andrade, A. C. M. A. d. Melo, and J. T. Cariolano, "Using ant's behavior-based simulation model AntWeb to improve website organization," in *SPIE's Aerospace/Defense Sensing and Controls Symposium:Data Mining*, Orlando, USA, 2002, pp. 229-240.
- [۷] C.-C. Lin, "Optimal Web site reorganization considering information overload and search depth," *European Journal of Operational Research*, vol. 173, no. 3, pp. 839-848, September 2006.
- [۸] B. J. Oommen and E. V. d. Croix, "String taxonomy using learning automata," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 27, pp. 354-365, 1997.
- [۹] H. Beigy, "Intelligent channel assignment in cellular networks: a learning automata approach," Ph.D. dissertation, Computer engineering department, Amirkabir university of technology, Tehran, 2004.
- [۱۰] A. B. Hashemi, H. Beigy, and M. R. Meybodi, "Dynamic Call Admission Control for Cellular Mobile Networks," in *9th Annual Computerf Society of Iran Computer Conference*, Tehran, Iran, 2004, pp.440-446.
- [۱۱] A. B. Hashemi, H. Beigy, and M. R. Meybodi, "Two Trashold Call Admission Control in Cellular Mobile Networks," in *12th Electrical Engineering Conference*, Mashhad, Iran, 2004, pp.37-42.
- [۱۲] A. B. Hashemi, H. Beigy, and M. R. Meybodi, "Dynamic Channel Assignment in Cellular Mobile Networks Using Learning Automata," in *12th Electrical Engineering Conference*, Mashhad, Iran, 2004.
- [۱۳] S .A. Myrer and M. G. Olsen, "Incremental Web Crawling as a Competitive Game of Learning Automata," M.S. thesis, Agder University College, 2005.
- [۱۴] A. B. Hashemi and M. R. Meybodi, "Web Usage Mining using Distributed Learning Automata," in *12th International CSI Computer Conference (CSICC'07)*, Tehran, Iran, 2006.
- [۱۵] M. A. L. Thathachar and R. H. Bhaskar, "Learning automata with changing number of actions," *IEEE Transactions on system, man and cybernetics*, vol. 17, no. 6, pp. 1095-1100, 1987.
- [۱۶] H. Beigy and M. R. Meybodi, "Utilizing Distributed Learning Automata to Solve Stochastic Shortest Path Problem," *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, vol. 14, no. 5, pp. 591-617, 2006.
- [۱۷] A. Asuncion and D. J. Newman" ,MSNBC Dataset," in *UCI Machine Learning Repository*: University of California, Irvine, School of Information and Computer Sciences, 2007.
- [۱۸] A. Asuncion and D. J. Newman, "Microsoft anonymous dataset," in *UCI Machine Learning Repository*: University of California, Irvine, School of Information and Computer Sciences, 2007.
- [۱۹] J. Liu, S. Zhang, and J. Yang, "Characterizing Web Usage Regularities with Information Foraging Agents," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 5, pp. 566-584.

## زیرنویس‌ها

---

<sup>۱</sup> Log file

<sup>۲</sup> Stationary

<sup>۳</sup> Non-Stationary

<sup>۴</sup> Learning automata with changing number of actions

<sup>۵</sup> Bias