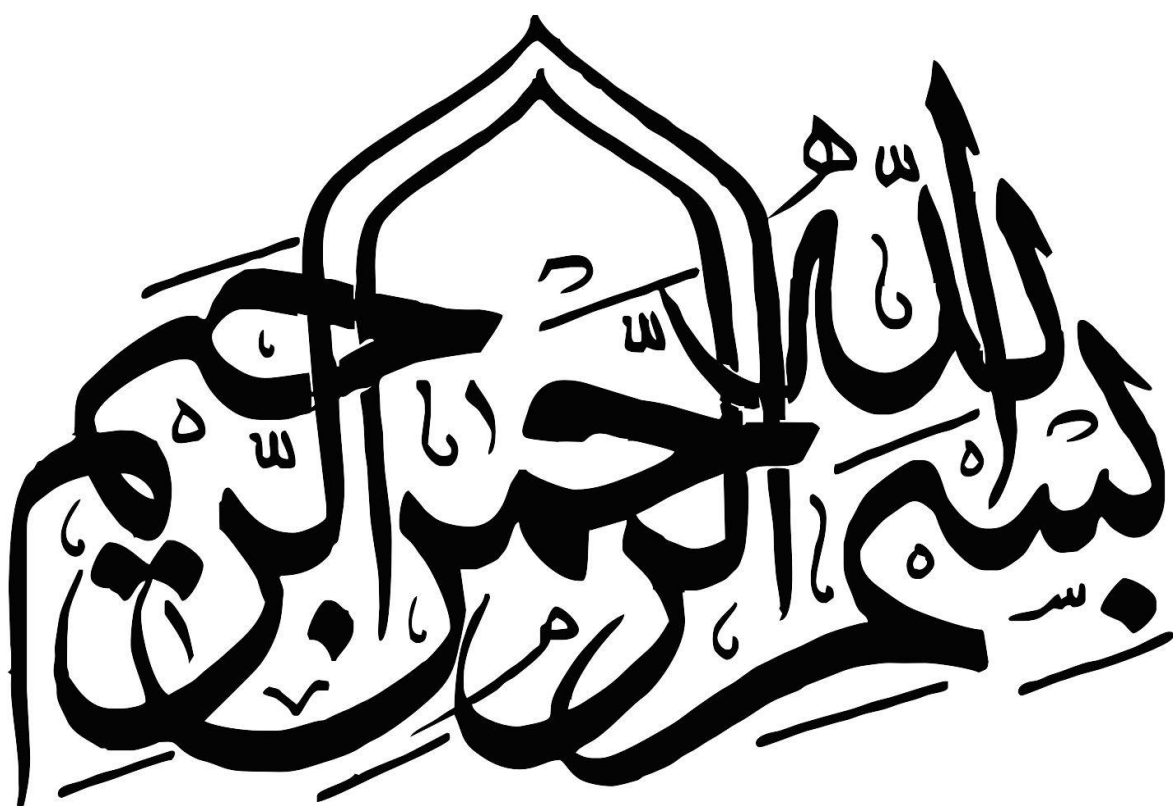






وزارت علوم تحقیقات و فناوری

دانشگاه فنی و حرفه ای

آموزشکده فنی و حرفه ای شماره یک تبریز

گروه کامپیوتر

عنوان پروژه:

طراحی موتور جستجوگر در اینترنت اشیاء و سئو در آن

پایان نامه جهت اخذ مدرک کارشناسی

گردآورنده: هیوا اعزامی

استاد راهنما:

مهندس منصور یوسفی نیا

مهندسی نرم افزار

زمستان 1396



وزارت علوم تحقیقات و فناوری

دانشگاه فنی و حرفه ای

آموزشکده فنی و حرفه ای شماره یک تبریز

گروه کامپیوتر

پایان نامه کارشناسی ناپیوسته برق - قدرت/ کامپیوتر هیوا اعزامی تحت عنوان طراحی مثنی جستنر در اینترنت اشیاء و سئو در آن

در تاریخ 96/12/10 مورد تایید و تصویب قرار گرفت و نمره پایان نامه به عدد و به

حروف اعلام می گردد.

مهندسی کامپیوتر

استاد راهنمای پایان نامه

از استاد گرامیم جناب آقای مهندس منصور یوسفی نیا بسیار سپاسگزارم چرا که بدون راهنماییهای ایشان تامین این پایان نامه بسیار مشکل می نمود.

از دوستان عزیزم وحید لطفی و سینا عسگری به دلیل یاری ها و راهنمایی های بی چشم داشت ایشان که بسیاری از سختی ها را برایم آسانتر نمودند،

و بعد از مدتها، پس از پیمودن راه های فراوان که با حضور شیرین اساتید عزیزم، با راهنمایی ها و دغده های فراوانشان و شیطنت های زیبای آن دوران، نگاه های پدر مادرم، با چشم های پر از برق شوق که خستگی های این راه را به امید و روشنی راه تبدیل کرده و امیدوارم بتوانم در آینده ی نزدیک جوابگوی این همه محبت آنها باشم.

فهرست مطالب

فصل اول: نیاز بشریت به اینترنت اشیا

10	چالش های روبرو در طراحی موتور جستجوگر در اینترنت اشیا.....	1-1
13	فصل دوم: اینترنت اشیا چیست؟	
13	موتور های جستجوگر چیست؟	1-2
13	تاثیر اینترنت اشیا بر زندگی انسان.....	2-2
13	رفتار عادی موتور های جستجوگر برای ایندکس سایت و خزنده ها	3-2
	فصل سوم: پیش نیاز های راه اندازی موتور جستجوگر در اینترنت اشیا.....	
15	بلاک چین Block Chain در اینترنت اشیا.....	1-3
16	کاربرد بلاک چین در اینترنت اشیا.....	2-3
19	پایگاه داده های سیار Mobile Database.....	3-3
20	محلی سازی در پایگاه داده های سیار.....	4-3
20	محلی سازی با پایگاه داده های سیار.....	4-3
21	معماری پایگاه داد های سیار.....	5-3
27	محلی سازی موتور جستجوی IoT با mobile database	6-3
22	Cloud در اینترنت اشیا	7-3
27	مشخصات اینترنت اشیا بر مبنای محاسبات ابری.....	8-3

فصل چهارم: تعریف موتور جستجو گر و طراحی آن

28	نکات طراحی موتور جستجو.....	1-4
28	معیار های طراحی موتور جستجو گر.....	2-4
29	آمار استفاده از موتور های جستجو در کشور های مختلف.....	3-4
32	اجزای موتور جستجو.....	4-4
32	هوش مصنوعی در پیاده سازی موتور جستجوگر IoT.....	5-4
33	هوش جمعی در موتور های جستجوگر.....	6-4
34	الگوریتم کلونی زنبور عسل در موتور های جستجوگر.....	7-4
35	ویژگی های خزنده خوب در موتور جستجوگر اینترنت اشیا.....	8-4

فصل پنجم: سئو یا Search Engine Optimization

38	تاریخچه سئو.....	1-5
39	اهداف سئو.....	2-5

41	انواع سئو.....	3-5
42	تکنیک های بهبود رتبه دستگاه در نتایج جستجو ها.....	4-5
	فصل ششم: الگوریتم های سئو برای داشتن بهترین نتایج	
46	Local Seo سئوی محلی.....	1-6
47	الگوریتم کبوتر.....	2-6
48	نتیجه گیری.....	
49	آینده اینترنت اشیا.....	
51	منابع.....	

فصل اول نیاز بشریت به اینترنت اشیا

فصل اول: نیاز بشریت به اینترنت اشیاء

همیشه ما برای یافتن جواب سوالات خود در گوگل سرچ میکنیم (Googling) چون پیدا کردن سایت مورد نیاز در بین میلیارد ها صفحه ایندکس شده در محیط اینترنت کار بسیار دشواری میباشد. این طرح در اوایل که تعداد سایت ها زیاد نبود به عنوان یک نیاز شناخته نشده بود اما بعد از مدتی به عنوان یک نیاز جدی برای کاربران مطرح شد که بعدا موتور های جستجوگر متنوع همانند گوگل و بینگ و یاهو و... در این حوزه پدیدار شدند.

امروزه که تعداد دستگاه های که در محیط اینترنت با اسم اینترنت اشیاء فعالیت میکنند بسیار زیاد شده است و احساس نیاز به یک موتور جستجوگر در آن احساس میشود اما راه اندازی یک موتور جستجو در IoT کار ساده ای نیست و چالش های زیادی را در برابر داریم.

تا کنون شما هم حتما چندین بار در گوگل عباراتی مثل "دانلود آهنگ" یا "بهترین غذاخوری های تهران" و... را شاید سرچ کرده باشید که با تعداد زیادی سایت در نتایج روبرو خواهید شد اما تصور نکنید که یخچال منزل شما اینبار به جای یک کاربر انسان، جستجوی خود را در موتور جستجو انجام دهد، یا شرایطی که کاربر انسان با سرچ در موتور جستجوگر به یخچال منزل شما دسترسی پیدا میکند.



شکل 1-1: موتور جستجوی اینترنت اشیاء

1-1 چالش های روبرو در اینترنت اشیاء

یکی از بزرگترین دغدغه هایی که پیش روی موتورهای جستجوگر قرار داد، آفلاین شدن دستگاه هامیباشد، همانطور که میدانید برخی از دستگاه هایی که به اینترنت اشیاء متصل هستند، از باتری استفاده میکنند و بعد از تمام شدن باتری آن ها دیگر امکان دسترسی به آن ها نیست و در شرایطی موتور جستجوگر این دستگاه را در نتایج نمایش دهد و دستگاه آفلاین باشد امکان ارتباط وجود ندارد و عملاً موتور جستجوگر ما مناسب نخواهد بود.

در محیط اینترنت صفحات توسط یکسری رباط خزنده یا Crawler شناسایی میشوند و از محتوای آن صفحه یک کپی گرفته میشود و در دیتابیس آن قرار میگیرد و از همان لحظه خزیدن بررسی داده ها انجام میشود اما در محیط IoT ما شبکه های نیمه بسته یا همان semi-closed network را داریم که اصلاً امکان دسترسی به آن شبکه وجود ندارد.

عمده ترین مشکلی که در ساخت سرچ اینجین IoT وجود دارد رجیستر یا ثبت کردن اولیه دستگاه در موتور جستجوگر میباشد. گوگل با استفاده از لینک هایی که وجود دارد معمولا در مدت زمان کمتر از 2 هفته آن سایت را ایندکس میکند و در نتایج خود نمایش میدهد. اما در اینترنت اشیاء این لینک و ارتباط وجود ندارد همچنین تاکنون چندین موتور جستجو تا کنون پیاده سازی شده اند که در همه آن ها رجیستر کردن سایت در اولین بار باید توسط خود کاربر و manual باشد اما یکی از چالش های روبرو در پیاده سازی موتور جستجوگر همین رجیستر شدن اتوماتیک میباشد. که برای رفع این مشکل نیاز به پیاده سازی یک مرکز NIC میباشد.

فصل دوم

اینترنت اشیاء چیست؟

فصل دوم: اینترنت اشیا چیست؟ اینترنت اشیا (Internet of Things (IoT)) مفهومی جدید در دنیای فناوری و ارتباطات است. به صورت خلاصه "اینترنت اشیا" فناوری مدرنی است که در آن برای هر موجودی (انسان، حیوان و یا اشیا) قابلیت ارسال داده از طریق شبکه‌های ارتباطی، اعم از اینترنت یا اینترنت، فراهم می‌گردد.

2-1 موتور های جستجوگر چیست؟

امروزه که تعداد زیادی از سایت ها فعالیت میکنند، و نیاز های کاربران هم بسیار گسترده تر شده و جستجو های حرفه ای تر هم انجام میشود، اما دسترسی به این محتوا و سایت ها اگر توسط خود کاربر انجام شود، ممکن است که کار بسیار دشواری باشد، اما برای رفع این مشکل از ابزار هایی به نام موتور جستجوگر استفاده میشود. این ابزار ها یک واسطه بین کاربر و سایت مقصد میباشد که با توجه به کوئری و سرچی که کاربران انجام داده اند مرتبط ترین سایت ها را برای آن ها نمایش میدهند. همچنین تعداد زیادی موتور جستجوگر وجود دارد، که هر کدام برای هدفی مشخص میباشد مثلاً برای جستجو های عمومی معمولاً کاربران از گوگل استفاده میکنند، یا برای جستجو در بین مقالات علمی و آرشیو های دولتی از موتور جستجوگر Wolframalpha استفاده میشود.

2-2 تاثیر اینترنت اشیا بر زندگی انسان

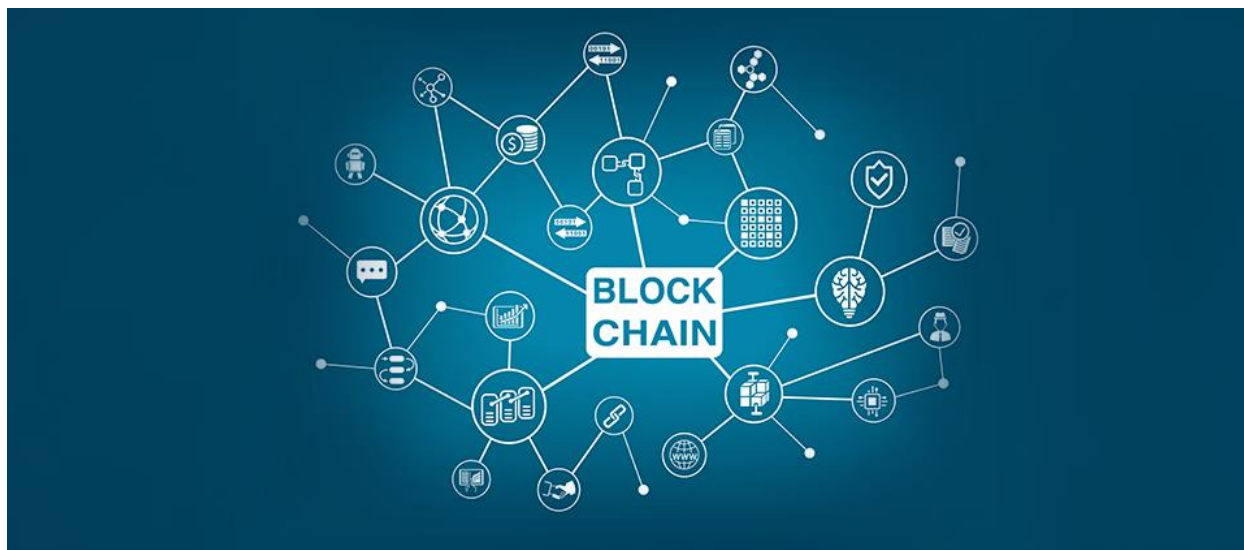
با توجه به رشد همه روزه این تکنولوژی و زیاد شدن تعداد دیوایس هایی که از این تکنولوژی استفاده میکنند و همه این تکنولوژی ها در زندگی روز مره ما نقش دارند، در رانندگی، میز صبحانه، خرید هایی که انجام میدهیم، و محیط کاری ما و تربیت فرزندان و ... یکی از دلایل رشد با سرعت بسیار بالای اینترنت اشیا تعبیه سازی شدن آن بر روی زندگی انسان ها میباشد.

2-3 رفتار عادی موتور های جستجوگر برای ایندکس سایت و خزنده ها

در حالت عادی موتور های جستجوگر با استفاده از کنسول وبمستر که خود ارائه میدهند سایت ها را پیدا میکنند همچنین از طریق لینک هایی که از طرف سایت های دیگر به سایت مورد نظر داده میشود موتور های جستجوگر این لینک ها را پیگیری میکنند و دنبال این لینک ها میروند و محتوای همه سایت هایی که در لینک ها قرار گرفته اند را ایندکس میکنند.

فصل سوم پیش نیاز های راه اندازی موتور جستجوگر در اینترنت اشیاء

با توجه به توضیح بالا دامنه اینترنت اشیا صرفاً محدود به دستگاه های دوروبر ما نیست، دربرگیرنده، انسان ها و حیوانات هم میشود بخاطر همین تعداد دستگاه ها بسیار بیشتر از محیط وب میباشد و در این میان است که تعاریفی همانند **Mobile Database** یا همان پایگاه داده های سیار و **Block Chain** مطرح میشود.



شکل 3-1 بلاک چین

3-1 Block Chain در اینترنت اشیاء

تا کنون بلاک چین بخاطر استفاده آن در بیت کوین دارای مثال های بانکی بود ، اما این تکنولوژی محدود به بیت کوین و سیستم های بانکی نیست و میتوان از آن در موتور های جستجو نیز استفاده کرد.

بلاک چین یک بستر توزیع تراکنش است که در آن هر تراکنش به دیگر تراکنش های اطراف آن از طریق رمزگذاری متصل است.

برای هک کردن آن، شما باید تمام تراکنش های وابسته به آن و تمام تراکنش های رونوشت شده را هک کنید (که در حقیقت یعنی غیر ممکن است).

بلاک چین دارایی های دیجیتالی شما را زیر نظر دارد تا فقط از طریق احراز هویت شناسایی شوند و کسی بدون اجازه نتواند کپی هایی از آن تهیه کند. چنین موضوعی سبب می شود که برای واحد پول چندان ارزشمند نباشد اما از طریق آن می توان بسیاری از مسائل دیجیتالی را حفاظت کرد. برای مثال، در آینده سند خانه شما دیگر یک برگه کاغذ نیست و یا کسی نمی تواند یک خانه را به چند نفر بفروشد و یا حتی رای گیری ها از این طریق کاملاً بدون خطا انجام می شوند.

فناوری بلاک چین، پیوند گمشده‌ای برای حل مسائل مربوط به :

- مقیاس پذیری
- حفظ حریم خصوصی
- قابلیت اطمینان در اینترنت است



شکل 2-3: شمای از ترکیب اینترنت اشیا و بلاک چین

3-2 کاربرد بلاک چین و اینترنت اشیا

طبق گفته سیسکو، 50 میلیارد دستگاه تا سال 2020 به اینترنت متصل می‌شوند. با این تعداد دستگاه متصل که همگی دستورات ارسال، دریافت و پردازش مانند روشن شدن، شماره گرفتن و حرکت کردن می‌فرستند، درستی داده‌ها در طول مسیر ممکن است هزینه‌های غیرقابل پیش‌بینی به بار آورد. مسائل دیگری مانند این که چگونه می‌توانیم به طور دقیق میلیاردها دستگاه متصل شده را ردیابی و مدیریت کنیم، چگونه داده‌های بزرگی که این دستگاه‌ها تولید می‌کنند، ذخیره کنیم و این که چگونه تمام این کارها را مطمئن و امن انجام دهیم، چیزهایی است که باید قبل از پذیرش مصرف کنندگان اینترنت اشیا حل شود.

شاید فناوری بلاک چین بتواند روزه امید برای صنعت IoT باشد. این فناوری می‌تواند برای ردیابی میلیاردها دستگاه متصل استفاده شود، معاملات را پردازش کرده و هماهنگی بین دستگاه‌ها را ممکن کند. این رویکرد غیرمتمرکز، احتمال از کار افتادن شبکه را از بین می‌برد و یک اکوسیستم مقاوم‌تر برای دستگاه‌ها ایجاد می‌کند. همچنین الگوریتم رمزنگاری استفاده شده توسط بلاک چین، داده‌های مشتریان را خصوصی‌تر می‌کند.

موارد کاربرد بلاک چین اینترنت اشیا

- خودروهای متصل
- لوازم هوشمند
- سنسورهای زنجیره تامین
- استفاده برای رجستر کردن دستگاه ها برای موتور های جستجوگر

3-4 محلی سازی در پایگاه داده های سیار

نکته: با توجه به تعریف dns که وظیفه آن تبدیل آدرس ip به اسم میباشد (دقیقا همانند دفترچه تلفن که از روی اسم کاربران شماره آن ها را پیدا میکنیم) امروزه تعداد زیادی دامنه اینترنتی ثبت شده اند با پسوند های مختلف، به نحوری که اگر شما قصد ثبت کردن یک دامنه جدید را داشته باشید به احتمال زیاد قبل از شما اشخاص دیگری آن دامنه را ثبت کرده باشند. و با در نظر گرفتن تعریف MD یا پایگاه داده سیار میتوان تا حد زیادی این دامنه گسترده را محدود تر کرد.



شکل 3-3: پایگاه داده های سیار

یکی از روش های مدیریت پایگاه های توزیع یافته Client\Server است. تمام بانکهای اطلاعاتی روی سرور ذخیره می شود و تمام پردازش های لازم نیز در سرور انجام می شود و تنها نتیجه پردازش ها به کلاینت فرستاده می شود.

اما پیشرفت های اخیر در تکنولوژی قابل حمل و بی سیم منجر به تولید کامپیوترهای سیار برای ارتباط داده ها و فرآیندها شده است که این کامپیوترها نیاز به ارتباطات بی سیم دارند. این نوع ارتباط به مشتری Client اجازه می دهد که کم و بیش به داده ها دسترسی داشته باشد و در اینجاست که ما از مدیریت mobile database استفاده می کنیم، اما همان گونه که می دانید در هر حال یک مسئله سخت افزاری و نرم افزاری در این نوع مسائل پیش می آید که باید حل شود.

بعضی از مسائل نرم افزاری که به مدیریت داده ها، اجرا و پوشش داده های پایه و مربوط می شود عمدتاً به خاطر پیوند متناوب و محدود کننده ی یک ارتباط بی سیم در این نوع کامپیوترها مشکل تر است. در این شبکه ها انتقال داده ها در یک واحد به نام Pack صورت می پذیرد که به طور معمول سرویس های دیگر هم استفاده می شود اما این روش mobile data بر روی پایگاه داده های توزیع شده اجرا می شود. این پایگاه داده ها مجموعه ای از داده های ذخیره شده است که منطقاً به یک سیستم تعلق دارند. ولی در کامپیوترهای مختلفی یک یا بیش از یک شبکه توزیع شده اند.

می توان گفت که در این معماری تعدادی پایگاه داده ها روی کامپیوترهای مختلف داریم که از نظر کاربران پایگاه واحد هستند. در معماری mobile data base یک یا بیش از یک کامپیوتر متوسط یا بزرگ نقش خدمتگزار Server پایگاه داده ها را ایفا می کند. این کامپیوتر را خدمتگزار میزبان Hostserver می نامند. هر کاربر کامپیوتر همراه خود را حمل می کند. در این کامپیوتر داده های عملیاتی و برنامه ی کاربردی مورد نیازش ذخیره شده اند.

پایگاه داده های ذخیره شده در کامپیوتر خدمتگزار میزبان حاوی داده های پویا Dynamic Data است. بدین ترتیب کاربر می تواند از هر جا با زیر سیستم خدمتگزار پایگاهی مرتبط بوده و پردازش های مورد نظرش را انجام دهد.

پردازش سیار برای افرادی که حتی در ضمن حرکت نیاز به اطلاعات دارند در این دنیایی که اطلاعات به سرعت در حال تغییر است به طور فزاینده در حال افزایش است. این مقاله به ارائه مفاهیم و مسائل اساسی مربوط به پایگاه داده های سیار می پردازد.

چرا پایگاه داده های سیار ؟

در گذشته پایگاه های داده بزرگ در مقیاس تجاری به عنوان سیستم های پایگاه داده مرکزی توسعه می یافتند. اما این روند با بیشتر و بیشتر شدن برنامه های کاربردی توزیع شده تغییر نمود. برنامه های کاربردی

توزیع شده معمولا درگیر پایگاه داده ها و شبکه های قدرتمند مرکزی هستند. همچنین، روند تکنولوژی جدید بخاطر فناوری های ظهور یافته مانند:

پیشرفت سریع فناوری های ارتباطی بی سیم و فن آوری کوچک سازی کامپیوتر کاربران را قادر به استفاده از منابع کامپیوتری در هر کجا از شبکه های کامپیوتری می کند. برای مثال ، شما حتی می توانید با اینترنت خود از هواپیما ارتباط برقرار کنید. بانک اطلاعاتی سیار پایگاه داده هایی هستند که به برنامه های کاربردی پایگاه داده اجازه توسعه و به کارگیری برای دستگاه های دستی را می دهند و در نتیجه برنامه های کاربردی مبتنی بر پایگاه داده رابطه ای قادر به بکارگیری توسط کاربران سیار می شوند. فن آوری پایگاه داده به کارمندان اجازه می دهد با استفاده از تجهیزات دستی به شبکه های شرکت خود لینک شده و به دریافت اطلاعات و کار آفلاین بپردازند و سپس با اتصال مجدد به شبکه دیگر با بانک اطلاعاتی شرکت همگام شوند. به عنوان مثال ، با یک پایگاه داده های سیار تعبیه شده در دستگاه دستی ، کارگر تحویل بسته می تواند مجموعه امضا ها بعد از هر بار تحویل را جمع آوری نموده و در پایان روز اطلاعات را به بانک اطلاعاتی شرکت ارسال کند.

3-3 پایگاه داده های سیار Mobile DataBase

پایگاه داده ای است که می تواند توسط یک دستگاه محاسباتی سیار از طریق شبکه سیار متصل شود بطوری که کلاینت و سرور دارای ارتباطات بی سیم هستند. همچنین یک حافظه نهان برای حفظ داده ها و تراکنشها هنگام قطع اتصال نگهداری می شود.. بانک اطلاعاتی ساختاری است برای سازماندهی اطلاعات، این اطلاعات ممکن است لیستی از مخاطبین ، قیمت اطلاعات، مسافت طی شده و ... باشد.

استفاده از لپ تاپ ها ، تلفن همراه و رایانه های جیبی در حال افزایش است و به احتمال زیاد در آینده با ظهور فزاینده برنامه های کاربردی سیستمهای سیار بیشتر نیز خواهد شد. در حالی که تحلیلگران نمی توانند محبوب ترین و عامترین برنامه در آینده را بطور دقیق پیش بینی کنند ، اما این کاملا روشن است که درصد زیادی از برنامه ها نیازمند استفاده از پایگاه داده خواهند بود. بسیاری از برنامه های کاربردی آینده نیازمند توانایی دانلود اطلاعات از مخزن اطلاعات و انجام عملیات بر روی این اطلاعات حتی زمانی که خارج از محدوده موردنظر بوده و یا زمانی که اتصال قطع شده می باشند.

امروزه با ظهور پایگاه های داده سیار، کاربران می توانند با تلفن های هوشمند خود و یا رایانه جیبی با پایگاه های داده سیار به تبادل اطلاعات از راه دور بدون هیچ گونه نگرانی در مورد زمان یا فاصله بپردازند. پایگاه های

داده سیار به کارکنان اجازه ورود اطلاعات حتی در همگام پرواز را می دهد. اطلاعات می توانند کمی بعد با سرور پایگاه داده هماهنگ شوند.

3-4 نیازمندیهای پایگاه داده های سیار

- برخی از عواملی که نیاز است در پردازش سیار در نظر گرفته شود ، عبارتند از :
- کاربران سیار باید بتوانند بدون اتصال بی سیم (حال یا به دلیل ضعیف بودن ارتباط و یا حتی هنگام قطع ارتباط) فعالیت خود را انجام دهند.
- زمان کاربر با ارزش ترین عامل در اکثر برنامه های کاربردی کسب و کار است
- با توجه به هزینه زمان اتصال، کاهش این زمان تا حد امکان بایستی مدنظر باشد
- تعداد بایت یا پکت در انتقال نیز از فاکتورهایی است که باید در محاسبه مدنظر باشد
- برنامه ها باید تعامل قابل توجهی ارائه دهد.
- برنامه ها باید قادر به دسترسی به دستگاه های محلی / سخت افزار مانند پرینتر ، اسکنر بار کد ، یا واحد GPS برای نقشه برداری و یا سیستم های خودکار محل یاب خودرو) باشند.
- پهنای باند باید محفوظ باشد (حداقل پهنای باند مورد نیاز در شبکه های بی سیم چند مگابایت است).
- کاربر نیاز به دسترسی به تاریخچه داده ها ندارد، تنها داده های تغییر یافته اخیر مورد نیاز است.
- زمان به روز رسانی اطلاعات باید زمان مناسبی باشد(در دوران اوج و یا خارج از دوره های اوج)
- عمر محدود منبع تغذیه (باتری)
- تغییر توپولوژی شبکه

3-5 معماری سیستم mobile database

- در هر گونه معماری های سیار ، موارد ذیل بایستی در نظر گرفته شود :
- کاربران در محل جغرافیایی ثابت نیستند
- دستگاه های کامپیوتری سیار: کم قدرت ، کم هزینه ، قابل حمل
- شبکه های بی سیم

- محدودیت های کامپیوتری

بخش های mobile database : پایگاه داده سیار به طور معمول شامل 3 بخش است: میزبان ثابت ، واحدهای سیار و ایستگاه های پایه.

3-6 محلی سازی موتور جستجوی IoT با mobile database

در 2 مرحله انجام میشود یکی در مرحله پیاده سازی برای بهبود عملکرد که از mobile database استفاده شده است و دیگری در قسمت SERP یا همان صفحه نتایج جستجو که از الگوریتم هایی مبتنی بر هوش مصنوعی استفاده میشود نمونه استفاده گوگل از این الگوریتم ها Local Seo میباشد که در ادامه بصورت کامل توضیح داده شده است.



شکل 3-4 : رایانش ابری در اینترنت اشياء

Cloud 7-3 در اینترنت اشياء

استفاده از شبکه ای از سرورها بر روی اینترنت برای ذخیره، مدیریت و پردازش داده ها را محاسبات ابری میگویند. از آنجا که نیاز به محاسبات پیچیده روز به روز بیشتر شده و برای انجام آنها نیاز به تجهیزات سخت افزاری و نرم افزاری خاصی هست و نگهداری حجم زیادی از داده ها نیز بر روی کامپیوترهای خانگی امکانپذیر نیست، محاسبات ابری کارایی خود را روز به روز بیشتر نشان میدهد، استفاده از سیستمهای cloud نه تنها در بسیاری از مواقع به صرفه تر است بلکه حتی قابلیت هایی فراتر از یک سیستم محلی بسیار قوی برای شما ایجاد میکنند. یک نمونه ساده بحث ذخیره سازی و آرشیو است. فرض کنید شما در جایی کار میکنید که الزم است روزانه ده ها عکس و فیلم را ذخیره کنید و تمامی کاربران سیستم هم به آنها دسترسی داشته باشند. خب قطعاً بهترین راه قرار دادن آنها در سرور است به جای اینکه آنها را بر روی تک تک کامپیوترها بریزید به خصوص که میزان مراجعه و دسترسی به آن فایلها برای هر فرد بسیار کم باشد. با قرار گرفتن همه آنها در یکجا و دسترسی دادن به کاربران داخل شبکه همه آنها میتوانند به راحتی و در هر زمان به آنها دسترسی داشته باشند و نیازی نیست کپی های متعددی بر روی همه سیستم ها موجود باشد، البته همانطور که میدانید سیستم های ابری مانند یک سرور FTP ساده تنها برای اشتراک فایل مورد استفاده قرار نمیگیرند و اساساً پیچیده تر هستند. با نگاهی به google play و apple store میتوان به اهمیت سیستم های ابری پی برد. نه تنها فایلهای

اپلیکیشن ها به صورت توزیع شده بر روی سرورهای ایل و گوگل قرار دارند بلکه انواع پروتکل های امنیتی و احراز هویت و سیستم های پرداخت درون برنامه ای و ... همگی بر روی یک بستر فراهم شده اند که در سراسر دنیا حتی با توجه به موقعیت جغرافیایی امکان فراهم کردن محتوای متفاوتی را ارائه میدهند. در واقع کار بر روی بسترهای محاسبات ابری با مباحث مربوط به اینترنت اشیا ارتباط بسیار تنگاتنگی دارند بهطوری که به معنای واقعی نمیتوان این دو مبحث را از هم جدا کرد.

مشخصات اینترنت اشیا بر مبنای محاسبات ابری

اولین مشخصه یک سیستم اینترنت اشیا بر مبنای محاسبات ابری این است که هر زمان به آن نیاز دارید آماده به خدمت است. منابع سیستم های توزیع شده در واقع سرویس های تحت وبی هستند که توسط کاربران و یا شما بدون نیاز به کمک و یا اجازه از طرف دیگر افراد قابل دسترسی میباشند، البته نیازی به توضیح نیست که ارتباط با اینترنت نیاز اساسی و پایه ای این سرویسها است. مشخصه دیگر این سیستم ها دسترسی گسترده شبکه است، به این معنا که روشهای ارتباطی متفاوتی برای دسترسی به آنها موجود است. میتوان از طریق دستگاه هایی که در یک شبکه خاص قرار دارند و به هر نحوی به اینترنت متصل هستند به منابع محاسبات ابری دسترسی پیدا کرد یعنی کافی است تنها یک نود از یک شبکه دسترسی به منابع محاسبات ابری داشته باشد. در واقع اینترنت اشیا و نوع کارکرد آن به نوعی امکان دسترسی و استفاده از منابع محاسبات ابری را آسانتر کرده به گونه ای که میتوان گفت اینترنت اشیا ضریب نفوذ اینترنت در اطراف ما را تقویت کرده است. از مشخصات دیگر این سیستمها میتوان به اشتراک منابع **pooling resource** اشاره کرد، به این معنا که سرویسهای محاسبات ابری می توانند برای همه آنهايي که آدرس منبع را میدهند قابل دسترس باشند. در دنیای اینترنت اشیا انسان ها می توانند به هر "شی" یک IP اختصاص دهند و به این ترتیب دسترسی آن به سیستمهای محاسبات ابری را فراهم نمایند. چهارمین خاصیت این سیستمها را میتوان انعطاف پذیری سریع **elasticity** نام برد به این معنا که شما به همان اندازه که منبع نیاز دارید دریافت میکنید و هر زمان که نیاز شما افزایش یافت به راحتی میتوانید منابع بیشتری در اختیار بگیرید. این خواص برای IOT اهمیت بسیار زیادی دارند چرا که با اضافه شدن اشیا به شبکه به راحتی میتوان منابع مورد نیاز آن را تامین کرد. در نهایت پنجمین خاصیت این سیستمها سرویسهای قابل اندازه گیری **measured service** است که اگر بخواهیم به زبان ساده تر آن را بیان کنیم میشود همان "هر چقدر بیشتر پول بدی بیشتر آش میخوری"! سیستمهای مبتنی بر محاسبات ابری میزان استفاده شده از فضا، توان محاسباتی، پهنای باند مصرفی، تعداد کاربرهای فعال در داخل شبکه ابری شما را اندازه گیری کرده و به نسبت میزان استفاده شما از شما هزینه دریافت میکنند. استفاده از سیستم های مبتنی بر محاسبات ابری بسیاری از کارها را برای توسعه دهندگان تسهیل می نماید.

در آینده نزدیک، باروند تمایل به تکنولوژی اینترنت اشیاء تعداد دستگاه هایی که به یکدیگر متصل می شوند، صدها برابر بیشتر از تعداد مردم خواهد بود. در سال 2012 میزان ترافیک داده هایی که فقط چندین خانوار ایجاد کردند برابر با کل ترافیکی بود که در سال 2008 در سراسر اینترنت ایجاد شد.

بهترین روش برای مدیریت و پردازش این حجم عظیم از داده ها پردازش ابری است. اینترنت اشیاء و پردازش ابری که اینگونه به یکدیگر ارتباط پیدا می کنند به عنوان ابر اشیاء نام گذاری می شوند ، با ابر اشیاء سرویس هایی که در ابر وجود دارند ارائه می شوند که یک دسترسی سراسری مکانی برای کاربران فراهم می آورد. همچنین سطوح کاربرد هم میتواند به راحتی دسترسی افزایش یابد.

با در نظر گرفتن موارد بالا در ادامه کاربرد IoT Cloud در هوش مصنوعی و پیاده سازی الگوریتم های موتور جستجوگر را بیان خواهیم کرد.

فصل چهارم تعریف موتور جستجوگر و طراحی آن

موتور جستجو چیست ؟

موتور جستجو یا پیمایشگر به طور عمومی به برنامه ای گفته می شود که کلمات کلیدی را در یک سند یا بانک اطلاعاتی جستجو می کند. در اینترنت به برنامه ای گفته می شود که کلمات کلیدی موجود در فایل ها و سندهای وب جهانی، گروه های خبری، منوهای گوفر و آرشیوهای FTP را جستجو می کند و به کاربران در جستجوی اطلاعات مورد نظر در میان انبوه اطلاعات موجود در شبکه کمک می کند. مانند

سایتهای Excite, Google, Yahoo, MSN

در موتور جستجو شما یک واژه ی کلیدی را وارد کرده و برنامه ی جستجو در بانک اطلاعاتی خود به جستجو پرداخته و سایت های مرتبط با موضوع شما را نمایش خواهند داد. البته امروزه بیشتر موتورهای جستجو دارای اطلاعات طبقه بندی شده مانند دایرکتوری ها می باشند و همچنین اکثر دایرکتوری ها قابلیت جستجو مانند موتورهای جستجو را دارند.

انواع موتور های جستجو:

موتورهای جستجو به دو دسته کلی تقسیم می شوند. موتورهای جستجوی پیمایشی و فهرستهای تکمیل دستی. هر کدام از آنها برای تکمیل فهرست خود از روش های متفاوتی استفاده می کنند البته لازم به ذکر است که گونه ای جدید از موتورهای جستجوگر تحت عنوان "ابر جستجوگر (Meta Search Engines)" نیز وجود دارد که در ادامه به توضیح هر یک از این موارد خواهیم پرداخت:

--موتورهای جستجوی پیمایشی:

موتورهای جستجوی پیمایشی (Crawler-Based Search Engines) مانند گوگل فهرست خود را بصورت خودکار تشکیل می دهند. آنها وب را پیمایش کرده، اطلاعاتی را ذخیره می کنند، سپس کاربران از میان این اطلاعات ذخیره شده، آنچه را که می خواهند جستجو می کنند. اگر شما در صفحه وب خود تغییراتی را اعمال نمایید، موتورهای جستجوی پیمایشی آنها را به طور خودکار می یابند و سپس این تغییرات در فهرستها اعمال خواهد شد. عنوان، متن و دیگر عناصر صفحه، همگی در این فهرست قرار خواهند گرفت.

--فهرست تکمیل دستی:

فهرستهای تکمیل دستی یا (Human-Powered Directories) مانند فهرست بازی (Open

Directory) مانند Dmoz وابسته به کاربرانی است که آن را تکمیل می کنند. شما صفحه مورد نظر را به همراه توضیحی کوتاه در فهرست ثبت می کنید یا این کار توسط ویراستارهایی که برای آن فهرست در نظر گرفته شده، انجام می شود. عمل جستجو در این حالت تنها بر روی توضیحات ثبت شده صورت می گیرد و در صورت تغییر روی صفحه وب، روی فهرست تغییری به وجود نخواهد آورد. چیزهایی که برای بهبود یک فهرست بندی در یک موتور جستجو مفید هستند، تأثیری بر بهبود فهرست بندی یک دایرکتوری ندارند. تنها استثناء این است که یک سایت خوب با پایگاه داده ای با محتوای خوب شانس بیشتری به نسبت یک سایت با پایگاه داده ضعیف دارد. البته در مورد جستجوگرهای مشهور از جمله گوگل و یاهو، یک مولفه دیگر هم برای بهبود فهرست بندی وجود دارد که کمک مالی است، یعنی وب گاههایی که مایل به بهبود مکان وب گاه خود در فهرست بندی هستند، می توانند با پرداخت پول به این جستجوگرها به هدف خویش برسند.

--موتورهای جستجوی ترکیبی با نتایج مختلط:

به موتورهای اطلاق می شود که هر دو حالت را در کنار هم نمایش می دهند. غالباً، یک موتور جستجوی ترکیبی در صورت نمایش نتیجه جستجو از هر یک از دسته های فوق، نتایج حاصل از دسته دیگر را هم مورد توجه قرار می دهد. مثلاً موتور جستجوی ام.اس.ان (MSN) بیشتر نتایج حاصل از فهرست های تکمیل دستی را نشان می دهد اما در کنار آن نیم نگاهی هم به نتایج حاصل از جستجوی پیمایشی دارد.

--ابر جستجوگرها:

این گونه جدید از موتورهای جستجوگر که قدمت چندانی نیز ندارند از ترکیب کردن نتایج حاصل از چندین موتور جستجوگر استفاده میکند. بدین معنی که این موتور عبارت مورد نظر شما را در چندین موتور جستجوگر جستجو میکند و نتایج آنها را با هم ترکیب کرده و یک نتیجه کلی به شما ارائه می دهد. به عنوان مثال موتور جستجوگر داگ پایل [1] از نتایج حاصل از موتورهای Google - Yahoo - MSN و ASK استفاده کرده و نتیجه حاصل را به شما می دهد

1-4 نکات طراحی موتور جستجو در اینترنت اشیا

+ اهمیت موتور های جستجو

+ اجزای موتور جستجو

+ انواع الگوریتم های رتبه بندی

+ موتور های جستجوی متن باز

+ موتور جستجو در lot

- مشکلات پیاده سازی موجود در IoT و وب فارسی با محدودیت های شبکه ای و فیلترینگ و تحریم های جهانی و...
- پیاده سازی خزنده یا کراولر در بین اشیاء
- شناسایی اشیاء فریب آمیز
- طراحی سیستم خبره فازی برای رتبه بندی درصد مناسب بودن اشیاء برای نمایش در نتایج جستجو
- الگوریتم های رتبه بندی در موتور جستجوی lot

2-4 معیار های طراحی موتور جستجو گر

بیش از 80 درصد از ترافیک اینترنت از طریق موتور های جستجو تامین میشود با این حال بی بهره بودن اینترنت اشیاء از این ویژگی نشان دهنده نیاز به ارتقاء در اینترنت اشیاء دارد.

همچنین با توجه به زیاد شدن اطلاعات و تعداد اشیاء متصل شده به اینترنت و با در نظر گرفتن Cloud در lot کاملاً نیاز به یک موتور جستجو خود را نمایان میکند.

از وظایف اصلی راهبران فناوری، تعبیه ساده تر نیاز ها میباشد، همانطور که تلفن های ثابت تعبیه شدند به گوشی های موبایل ما نیز باید همه تلاش خود را برای تعبیه سازی اینترنت اشیاء انجام دهیم. اما تعبیه سازی فقط راحتی در سیرکولاسیون و دسترسی پذیری نیست، باید از نظر فرهنگ، زبان و سوابق قبلی تعبیه سازی انجام شود و این موارد نیز در الگوریتم های سئو نیز باید در نظر گرفته شود.



شکل 1-4: معیارهای نمایش دستگاه ها در نتایج جستجو

عامل موفقیت هر موتور جستجو وجود قوانین و اصول مناسب در نمایش بهترین نتایج با در نظر گرفتن نکات بالا میباشد. با توجه به این توضیحات تاکنون موتور جستجو های فراوانی برای جستجو در **lot** تا کنون طراحی شده اند اما هیچکدام تا کنون موفق نشده اند و بصورت رسمی نیز شناخته نشده اند. حتی موتور های جستجوگر فارسی نیز طراحی شده اند. اما با در نظر گرفتن اینکه اگر موتور جستجو نتواند بهترین نتیجه را به کاربران ارائه بکند، کاربران دیگر از آن استفاده نمیکنند. یکی از علت های پیروزی گوگل بر یاهو نیز همین مورد بوده است. بهترین **SERP** یا همان **Search Engine Result Page** آن است که توسط یکی از دوستان شما پیشنهاد شود، اکنون اگر با استفاده از هوش مصنوعی استفاده کرد و داده ها و تاریخچه جستجو را آنالیز کرد میتوان نتیجه ای مطلوب تر را داشت.

3-4 آمار استفاده از موتور های جستجو در کشورهای مختلف

Country	Search Engines			
	Leader	Share	Runner Up	Share
Argentina	Google	95%	Bing	4%
Australia	Google	87%	Bing	3%
Brazil	Google	97%	Bing	2%
Canada	Google	78%	Bing	6%
Czech Republic	Seznam	45%	Google	45%
China	Baidu	76%	Google	22%
Denmark	Google	97%	Bing	2%
Egypt	Google	95%	Yahoo/Bing	5%
Finland	Google	95%	Other	5%
France	Google	92%	Bing	4%
Germany	Google	89%	t-online	3%
Hong Kong	Yahoo	N/A	Google	N/A
India	Google	95%	Other	5%
Indonesia	Google	92%	Yahoo	5%
Israel	Google	89%	Other	11%
Italy	Google	87%	Virgilio	5%
Japan	Yahoo Japan*	56%	Google	31%
Malaysia	Google	92%	Yahoo	5%
Mexico	Google	91%	Bing	7%
The Netherlands	Google	94%	Vinden	3%
New Zealand	Google	93%	Bing	2%
Philippines	Google	N/A	Yahoo	N/A
Poland	Google	97%	Other	3%
Russia	Yandex	62%	Google	26%
Saudi Arabia	Google	97%	Yahoo/Bing	3%
Slovakia	Google	99%	Yahoo/Bing	1%
South Korea	Naver	73%	Daum	18%
Spain	Google	96%	Yahoo/Bing	4%
Turkey	Google	N/A	Yandex	N/A
United Kingdom / UK	Google	94%	Bing	5%
United States	Google	72%	Yahoo	14%

شکل 2-4: موتور جستجوگرهای موفق در کشورهای مختلف

برای توضیحات دقیق تر در زیر در کشور های مختلف محبوب ترین موتور های جستجوگر لیست شده اند که علت موفقیت برخی از آن ها مثلا Baidu در چین تعبیه سازی و رعایت شرایط بالا بهتر از گوگل میباشد همچنین تحریم ها نیز بر روی موتور های جستجوگر موثر بوده اند مثلا در کشور کره جنوبی از موتور جستجوگر

Naver استفاده میشود اما حتی با سرچ کلمه کلیدی : info:Naver.com در گوگل متوجه خواهید شد که گوگل هیچ اطلاعاتی از این سایت را ثبت نکرده است

با استناد به آمار سایت Alexa که بصورت قطع و 100 درصد صحیح نمیشد اما تقریباً آمار هایی درست را میدهد و در بین سایت های رقیب بهترین سایت میباشد که نزدیک ترین نتیجه را ارائه میکند، سایت naver.com که در گوگل هیچ اطلاعات ثبت شده ای ندارد، حتی بسختی میتوان در بین سایت های دیگر هم توضیحات آن را پیدا کرد اما دارای آمار بازدید بسیار بالایی میباشد که جزو سایت های اول دنیا میباشد. نمونه گزارش الکسا:

How popular is naver.com?

Alexa Traffic Ranks

How is this site ranked relative to other sites?



Global Rank ?

124 ▼ 9

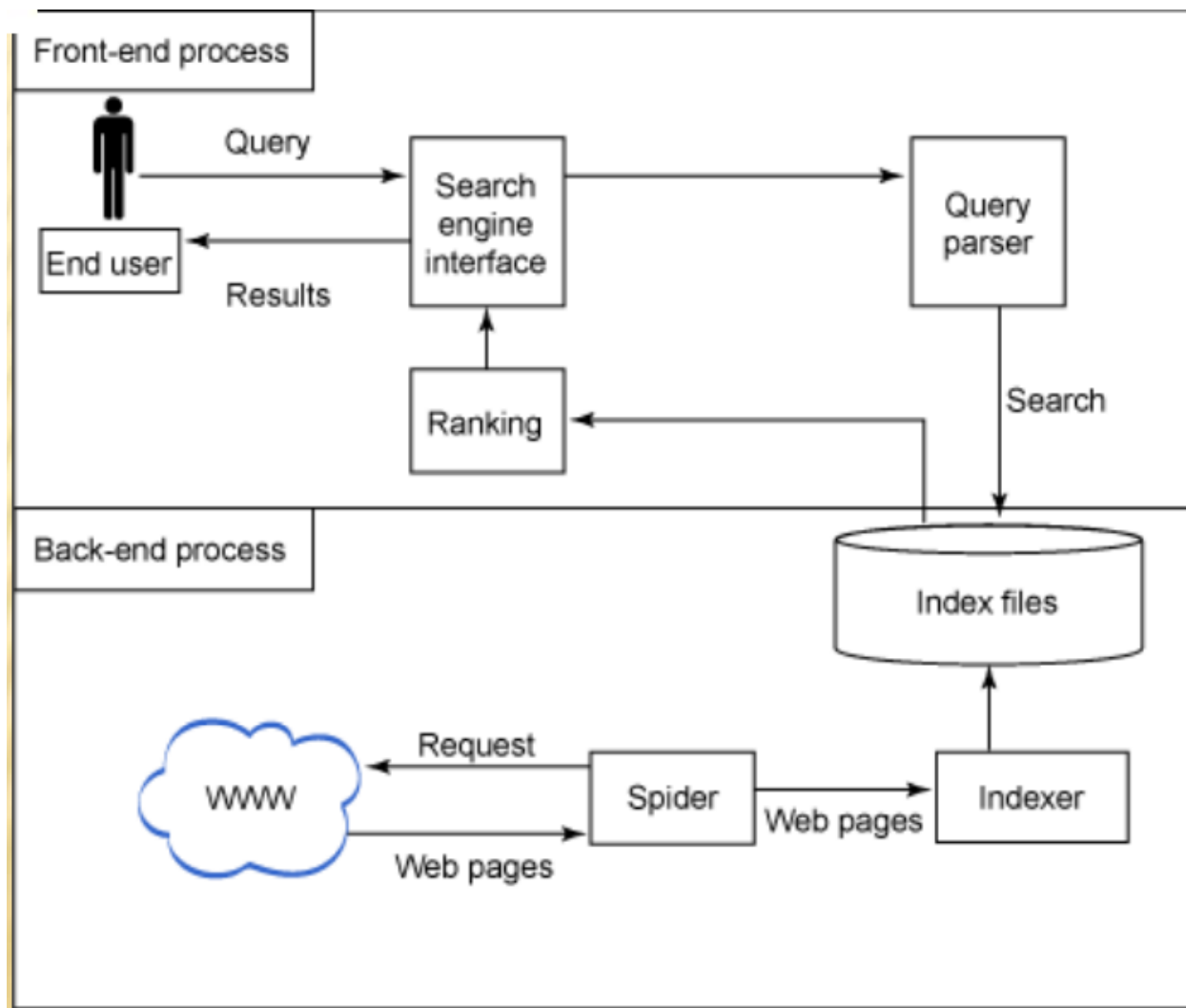
Rank in South Korea ?

2

رتبه دوم در کره جنوبی و 124 در کل دنیا، با توجه به نمودار مشخص میباشد که رتبه این سایت در یک دوره خاص از زمان کمتر از 60 دنیا بوده است و اخیراً نیز افت رتبه داشته است..

در استفاده از alexa باید یک نکته را در نظر داشت، باتوجه به اینکه اطلاعات و داده های این سایت توسط افزونه alexa toolbar گرفته میشود بخاطر همین اگر تعداد مخاطبان و کاربرانی که از این افزونه در مرورگر استفاده میکنند بیشتر باشد نتیجه نیز بهتر خواهد بود همچنین این شرکت با برخی از شرکت های ارائه مرورگر اینترنت نیز قرار داد دارد و اطلاعات خود را با یکدیگر به اشتراک میگذارند.

4-4 اجزای موتور جستجو



شکل 3-4: پروسه نمایش نتایج برای کلاینت

در شکل بالا از کاراکتر انسان استفاده شده است اما در اینترنت اشیاء میتواند انسان حیوان یا هر شی دیگری باشد.

5-4 هوش مصنوعی در پیاده سازی موتور جستجوگر IoT

با توجه به نیاز هایی که در بالا تعریف شد، اعم از فضای ابری ما نیاز به پیاده سازی الگوریتم هایی برای بهینه سازی داریم. برای اینکار الگوریتم های بهینه سازی متنوعی وجود دارند اما بعد از تجزیه و تحلیل انواع دستگاه ها (هوش ازدحامی یا هوش گروهی (Swarm Intelligence) و الگوریتم کلونی زنبور عسل ها را مطرح کرده ایم. در بین موتور های جستجو همانند گوگل که از هوش مصنوعی بهره میبرند از الگوریتم های دقیق تر و

بسیار پیچیده تری استفاده شده است اما مدیران این شرکت ها برای جلوگیری از سئوی کلاه سیاه معمولاً نحوه عملکرد الگوریتم های خود را منتشر نمیکنند و کسی از آن ها اطلاعی ندارد.

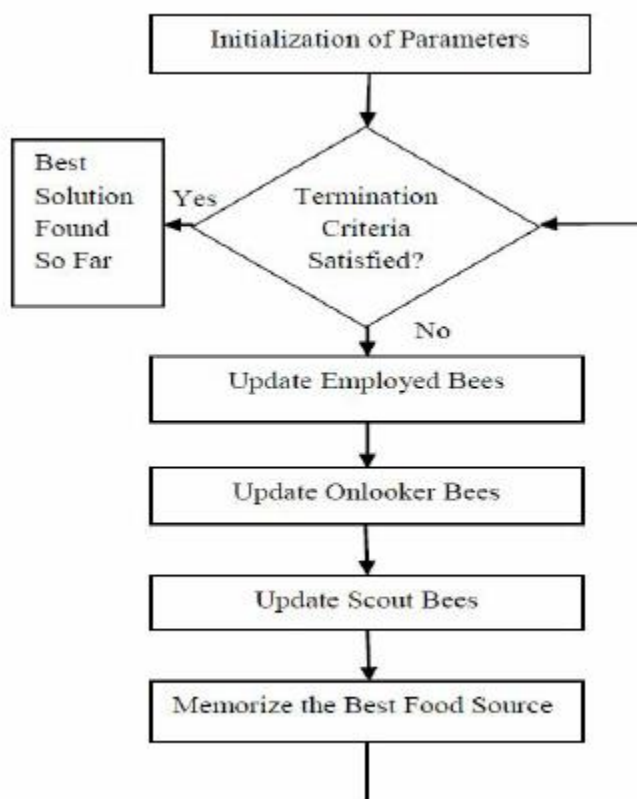
4-6 هوش جمعی یا Swarm Intelligence

فرض کنید شما و گروهی از دوستانتان به دنبال گنج می گردید. هر یک از اعضای گروه یک فلزیاب و یک بی سیم دارد که می تواند مکان و وضعیت کار خود را به همسایگان نزدیک خود اطلاع بدهد. بنابراین شما می دانید آیا همسایگانتان از شما به گنج نزدیکترند یا نه؟ پس اگر همسایه ای به گنج نزدیکتر بود شما می توانید به طرف او حرکت کنید. با چنین کاری شانس شما برای رسیدن به گنج بیشتر می شود و همچنین گنج زودتر از زمانی که شما تنها باشید، پیدا می شود.

این یک مثال ساده از رفتار جمعی یا Swarm behavior است که افراد برای رسیدن به یک هدف نهایی همکاری می کنند. این روش مؤثرتر از زمانی است که افراد جداگانه عمل کنند. Swarm را می توان به صورت مجموعه ای سازمان یافته از موجوداتی تعریف کرد که با یکدیگر همکاری می کنند. در کاربردهای محاسباتی Swarm intelligence از موجوداتی مانند مورچه ها، زنبورها، موریانه ها، دسته های ماهیان و دسته ی پرندگان الگو برداری می شود. در این نوع اجتماعات هر یک از موجودات ساختار نسبتاً ساده ای دارند ولی رفتار جمعی آنها بی نهایت پیچیده است. برای مثال در کولونی مورچه ها هر یک از مورچه ها یک کار ساده ی مخصوص را انجام می دهد ولی به طور جمعی عمل و رفتار مورچه ها، ساختن بهینه ی لایه، محافظت از ملکه و نوزادان، تمیز کردن لانه، یافتن بهترین منابع غذایی و بهینه سازی استراتژی حمله را تضمین می کند. رفتار کلی، یک Swarm به صورت غیر خطی از آمیزش رفتارهای تک تک اجتماع بدست می آید. یا به عبارتی یک رابطه ی بسیار پیچیده بین رفتار جمعی و رفتار فردی یک اجتماع وجود دارد. رفتار جمعی فقط وابسته به رفتار فردی افراد اجتماع نیست بلکه به چگونگی تعامل میان افراد نیز وابسته است. تعامل بین افراد، تجربه ی افراد درباره ی محیط را افزایش می دهد و موجب پیشرفت اجتماع می شود. ساختار اجتماعی Swarm بین افراد مجموعه کانالهای ارتباطی ایجاد می کند که طی آن افراد می توانند به تبادل تجربه های شخصی بپردازند، مدل سازی محاسباتی Swarm ها کاربردهای موفق و بسیاری را در پی داشته است مانند:

Function optimization, Finding optimal roots, scheduling, structural optimization, Image and data analysis

کاربردهای زیادی از مطالعه‌ی Swarm‌های مختلف وجود دارد. از این دسته می‌توان به کولونی مورچه‌ها Ant Colony و دسته‌ی پرندگان Bird Flocks اشاره نمود.



شکل 4-4 : پروسه انجام هوش جمعی در موتور جستجو

4-7 الگوریتم کلونی زنبور عسل در موتور های جستجوگر

الگوریتم کلونی زنبور عسل مانند سایر الگوریتم های هوش ازدحامی مرتبط بر رفتار تصادفی المان های آن است و برای حل مسائل بهینه سازی کاربرد دارد. بسیاری از الگوریتم های هوش ازدحامی با الهام گرفتن از طبیعت ایجاد شده اند مانند الگوریتم کلونی مورچگان، الگوریتم پرندگان، الگوریتم فاخته و الگوریتم کلونی زنبور عسل یا Artificial bee colony algorithm که به صورت مخفف BCO نامیده میشود (Bee Colony Optimization) .

برخی کاربردهای الگوریتم بهینه سازی زنبور عسل در علوم مهندسی به صورت زیر است:

- آموزش شبکه عصبی برای الگو شناسی

- زمان بندی کارها برای ماشین‌های تولیدی
- دسته‌بندی اطلاعات
- بهینه‌سازی طراحی اجزای مکانیکی
- بهینه‌سازی چند گانه
- میزان کردن کنترل کننده‌های منطق فازی برای ربات‌های ورزشکار

بسیاری از مسائل به روش‌های معمول ریاضی قابل حل نیستند و یا حل کردن آنها زمان بسیار زیادی را می‌طلبد. در این نوع از مسائل ما به دنبال پیدا کردن یک نقطه بهینه در مسئله هستیم که اصطلاحاً به آن نقطه، نقطه بهینه می‌گوییم. نقطه بهینه زمانی بدست می‌آید که ما کمترین خطا در مسئله را داشته باشیم. الگوریتم‌هایی تصادفی مانند الگوریتم ژنتیک و الگوریتم‌های تکاملی برای حل مسائل بهینه‌سازی استفاده می‌شوند.

یکی دیگر از روش‌های حل مسائل بهینه‌سازی الگوریتم‌های هوش ازدحامی است که الگوریتم زنبور عسل از جمله این الگوریتم‌ها است. الگوریتم زنبور (Bee Algorithm) یک الگوریتم گروهی مبتنی بر جستجو است که در سال ۲۰۰۵ میلادی ابداع شده است. این الگوریتم شبیه‌سازی رفتار جستجوی غذای گروه‌های زنبور عسل است. در نسخه ابتدایی این الگوریتم، الگوریتم نوعی از جستجوی محلی انجام می‌دهد که با جستجوی تصادفی کتره-۱ ترکیب شده و می‌تواند برای بهینه‌سازی ترکیبی یا بهینه‌سازی تابعی استفاده شود.

این الگوریتم نیز مانند سایر الگوریتم‌های هوش ازدحامی از دو روش اکتشاف و استخراج استفاده می‌کند. زنبورهای کارگر وظیفه استخراج و زنبورهای ناظر وظیفه اکتشاف را به عهده دارند. زنبورهای کارگر در اطراف یک منطقه (گل‌های پیدا شده یا منطقه‌ای که شامل جواب مسئله است) به دنبال جواب بهینه می‌گردند و زنبورهای ناظر با رفتار تصادفی به دنبال پیدا کردن مناطق جدید هستند (گل‌های جدید)

8-4 ویژگی‌های خزنده خوب در موتور جستجوگر اینترنت اشیاء

- ایندکس نکردن دستگاه‌های تکراری
- امکان اجرا در Threadهای موازی بسیار
- گیر نیفتادن در چاله دستگاه‌ها
- از نظر سرعت

○ دستگاه های بدرد نخور و اسپم

- بروزرسانی دستگاه ها برای بررسی وضعیت آنها
- جستجوی محلی
- نمایش بهترین نتیجه بر اساس درخواست کننده

فصل پنجم سئو یا

Search Engine Optimization



شکل 1-5: یافتن بهترین نتایج در جستجوگر

1-5 تاریخچه سئو

سئو برگرفته از عبارت Search Engine Optimization است که کلمات معادلی چون SEO یا بهینه سازی موتور جستجو دارد که همگی آنها به یک معنی می باشند و می توان آن را بصورت زیر تعریف نمود:

سئو (بهینه سازی موتور جستجو) به معنای فریب و یا دستکاری موتورهای جستجو با روش های غیر اخلاقی نیست. آن ها برای کمک کردن به بهبود رویت و ارتباط وب سایت ها در نتایج جستجوی طبیعی با هدف رسیدن به رتبه بندی بالا، ایجاد شده اند.

سئو را می توان مجموعه ای از روش ها برای تغییر استراتژیک وب سایت ها دانست. این فرآیند به بیان مهمترین عوامل مرتبط صفحه و افزایش اهمیت آن در صفحه نتایج جستجو، می پردازد.

سئو فرآیند ساده ای نیست که به راحتی پیاده سازی شود زیرا نیازمند دانش زیاد، علم پیش زمینه ای و صبر است. موتورهای جستجو با تغییر مداوم الگوریتم های رتبه بندی خود غیر قابل پیش بینی هستند؛ بنابراین وظیفه مهندسين سئو بروز نگه داشتن اطلاعات خود در این زمینه است.

با توجه به رشد و روزرسانی مداوم گوگل و استاندارد های مناسبی که دارد در توضیحات زیر قوانین سئو و بهینه سازی گوگل را در نظر گرفته ایم که قابلیت پیاده سازی در موتور جستجوگر اینترنت اشیاء را نیز دارد.

الگوریتم رتبه‌بندی گوگل دارای بیش از 200 معیار جهت رتبه‌بندی یک سایت است. از این رو مهندسين سئو می‌بایست مهم‌ترین این معیارها را جهت پیاده‌سازی موفق سئو بدانند.

با رشد مداوم اینترنت موتورهای جستجوی جدیدتری ظاهر می‌شدند که از الگوریتم‌های رتبه‌بندی پیچیده‌تری استفاده می‌کردند. این الگوریتم‌ها با معیار تراکم کلمه کلیدی (تعداد دفعاتی که یک کلمه کلیدی یا یک عبارت در یک صفحه استفاده شده تقسیم بر تعداد کل کلمات صفحه) و برچسب‌های شبیه key-words برای فهم محتویات وب سایت‌ها استفاده می‌کردند. به دنبال این رشد سئو برای بیشتر دیده شدن سایت‌ها فعالیت خود را شروع کرد. با توسعه رتبه‌بندی صفحات توسط موتورهای جستجو، مهندسين سئو به دنبال پیدا کردن راه‌هایی جهت تاثیرپذیری در رتبه بندی این موتورها چه با استفاده از روش‌های اخلاقی و چه غیر اخلاقی روی آوردند.

یک مقاله که توسط نیویورک تایمز در نوامبر سال 1996 چاپ شد راجع به توسعه‌دهندگان وب صحبت کرد که برای قرار گرفتن در بالای صفحه نتایج جستجو از روش‌های غیر اخلاقی سئو استفاده می‌کردند. این مقاله توضیح می‌دهد که چگونه توسعه‌دهندگان وب به راحتی یک سایت را به همراه کلمات کلیدی ویژه که معمولاً در پشت گرافیک سایت پنهان شده است، بارگذاری می‌کنند که در نتیجه موتورهای جستجو این کلمات کلیدی را شمارش کرده و رتبه بندی سایت را افزایش می‌دهند.

همان طور که گفته شد بدست آوردن رتبه‌بندی بالا برای کلمات کلیدی خاص در اوایل کار خیلی راحتی نسبت به امروز بوده است. با بکارگیری روش‌های غیر اخلاقی در آن زمان قرار گرفتن در بالای لیست نتایج جستجو تضمین شده بود. امروزه الگوریتم‌های موتورهای جستجو پیچیده‌تر شده‌اند و خیلی سخت است که مهندسين بهینه‌سازی بتوانند موتورهای جستجو را دستکاری کنند، البته این بازی موش و گربه بین مهندسين سئو و موتورهای جستجو امروزه نیز وجود دارد و ادامه پیدا خواهد کرد.

2-5 اهداف سئو چیست ؟

هدف سئو کمک به وب سایت‌ها در بدست آوردن محل بالا در نتایج جستجوی طبیعی و افزایش ارتباط بین صفحات و کلمات کلیدی وارد شده توسط کاربران است. نمایش دادن بیشترین صفحات مرتبط با پرس‌وجوی جستجوی وارد شده دارای مزایای زیادی هم برای کاربر و هم برای موتور جستجو است که در آن کاربر بیشترین نتایج مرتبط جستجو شده را پیدا می‌کند و موتور جستجو قابل اعتماد می‌شود چرا که الگوریتم آن بیشترین صفحات مرتبط را نشان می‌دهد.

بعد از اینکه سئو برای بهبود رتبه‌بندی سایت‌ها در نتایج جستجوی طبیعی کار خود را شروع کرد، این فرآیند نیازمند دانش و زمان برای پیاده‌سازی روش‌های خود شد. زمان، بزرگ‌ترین سرمایه‌گذاری اولیه برای بهینه شدن سایت و یا یک صفحه است.

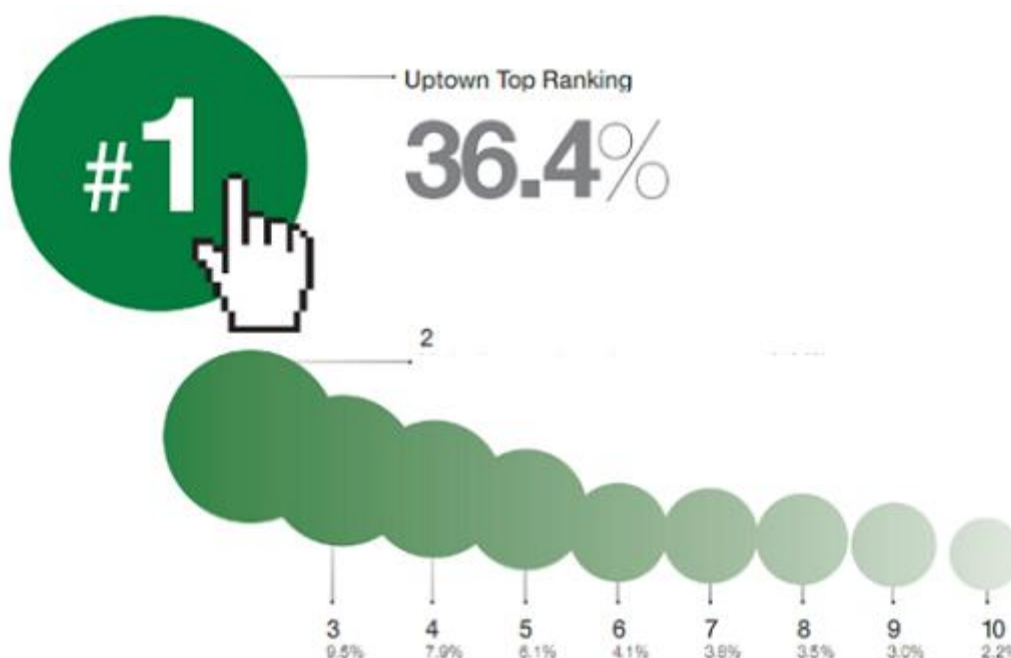
بسته به اینکه چگونه کلمات کلیدی رقابتی در بهینه سازی مورد استفاده قرار می گیرند، سئو هفته ها یا ماه ها طول می کشد تا به نتیجه برسد. اگر چه ممکن است نتایج بهینه سازی جستجو به سرعت دیده نشود ولی در دراز مدت مزایای سئو باعث افزایش رتبه بندی سایت و افزایش بازدیدکنندگان است. برای وب سایت ها خیلی مهم است که در صفحه اول نتایج جستجو گوگل به خصوص جزو سه تا نتیجه طبیعی اول که 58.4 درصد کلیک ها را شامل می شود، قرار گیرند.

این جای تعجب نیست که وب سایت ها با محصولات و خدماتی که بفروش می رسانند بر سر رسیدن به رتبه یک نتایج جستجو رقابت کنند. به عنوان مثال کلمه کلیدی رقابتی "بیمه خودکار" 1.5 میلیون جستجوی گوگل را در ماه به خود اختصاص می دهد، این بدین معنی است که 3 جایگاه اول 58.4 درصد بازدیدکنندگان و خود جایگاه اول 36.4 درصد بازدیدکنندگان را به خود اختصاص می دهد. این مسئله برای خیلی از سایت ها سودمند است که در بالای نتایج جستجو قرار گیرند به خصوص آنهایی که با افزایش ترافیک سایت به تعداد مشتریان اضافه می شود.

یک مطالعه دیگر نشان می دهد جستجو کنندگان خیلی دوست دارند روی پیوندهای طبیعی نتایج جستجو کلیک کنند؛ این مطالعه مشخص می کند 72.3 درصد کاربران گوگل روی پیوندهای طبیعی نتایج جستجو کلیک می کنند و تنها 27.3 درصد کاربران روی پیوندهای لیستهای پرداخت شده کلیک می کنند. این یکی دیگر از دلایلی است که شرکت ها تاثیر سئو را دانسته و حاضر هستند برای آن زمان و هزینه صرف کنند.

بطور کلی اهداف سئو را می توان بصورت زیر خلاصه نمود:

- در بازدید دستگاه یا اشیاء متصل به اینترنت تاثیر بسزایی دارد.
- باعث ایجاد برند می شود.
- باعث افزایش ترافیک دستگاه یا شی متصل به اینترنت خواهد شد.
- باعث شده دستگاه در صفحه نتایج موتور جستجو از رتبه بهتری برخوردار شد.
- باعث افزایش فروش محصولات خواهد شد.
- بهترین ابزار برای تبلیغات است.



شکل 2-5: نرخ بازدید از نتایج جستجوها بر اساس رتبه

3-5 انواع سئو

سئو درون صفحه چیست ؟ On Page SEO

سئو درون صفحه به کنترل مستقیم کدها و یا محتویات وب سایت شما اعم از متن، تصویر، پیوند و ... مرتبط می‌شود؛ به طور کلی هر چیزی که در سایت خود بارگذاری می‌کنید مربوط به **SEO On Page** می‌شود. این روش پایه تمام تلاش‌های بهینه‌سازی است زیرا جایی است که شما بیشترین کنترل را روی آن دارید. هر گونه تغییری را که شما در محتویات سایت خود ایجاد می‌نمایید در نتایج جستجو تاثیر گذار خواهد بود بنابراین خیلی مهم است که روش **SEO On Page** بدرستی قبل از روش **SEO Off Page** انجام شود.

سئو بیرون صفحه چیست ؟ Off Page SEO

بر خلاف روش **SEO On Page** که توسعه وب سایت با کنترل مستقیم بالایی انجام می‌شد، روش بهینه سازی بیرون صفحه شامل تمام ارتقاها و تغییراتی است که در خارج از وب سایت شما انجام می‌شود؛ این شامل تمام چیزهایی است که می‌تواند در بدست آوردن رتبه‌بندی بالای سایت شما کمک کند. **SEO Off Page** در درجه اول در بدست آوردن پیوند دهندگان (پیوندهایی که از سایت‌های دیگر به سایت شما اشاره می‌کنند) متمرکز می‌شود سپس با استفاده از روش‌هایی که در خارج از وب سایت انجام می‌شود به بهینه‌سازی موتور جستجو می‌پردازد.

SEO Off Page یک فرآیند نامحدود است که توسط مهندسين سئو در جهت بدست آوردن پیوند دهندگان، اداره می‌شود.

سئو کلاه - سفید چیست ؟ White Hat

در اصطلاحات فنی این حوزه، سئو کلاه - سفید به مجموعه تکنیک‌ها، روش‌ها و استراتژی‌های که از قوانین و راهنمایی‌های موتورهای جستجو تبعیت می‌کنند، اشاره دارد. که این به معنی فریب یا دستکاری موتورهای جستجو و یا کاربران به هیچ وجه نیست. این همچنین به سئو اخلاقی نیز یاد می‌شود زیرا دلیل آن این است که یا اجرای بلند مدت داشته و یا از دستورالعمل‌ها و سیاست‌های صحیح جهت بدست آوردن رتبه‌های بالا در نتایج موتورهای جستجو استفاده می‌کند.

سئو کلاه - سیاه چیست ؟ Black Hat

به عبارتی دیگر اصطلاح سئو کلاه - سیاه اشاره به مجموعه‌ای از دستورات دارد که به منظور فریب، بازی و دستکاری الگوریتم‌های موتورهای جستجو با هدف بدست آوردن رتبه‌بندی بالا، بکار می‌رود. این مهم است که بدانید این روش‌ها غیر اخلاقی بوده و توسط هیچ موتور جستجویی تایید نمی‌شود و بکارگیری این روش‌ها خطر خارج شدن از فهرست پایگاه داده را برای همیشه در بر دارد.

انجام تاکتیک‌های سئو کلاه - سیاه ممکن است که باعث افزایش رشد ناگهانی در نتایج جستجو شود ولی می‌تواند منجر به پرداخت جریمه و یا حذف کامل از فهرست پایگاه داده موتور جستجو شود.

4-5 تکنیک های بهبود رتبه دستگاه در نتایج جستجو ها

- بهبود ساختار دستگاه و دیوایس
- استراتژی کلمات کلیدی در رجیستر کردن دستگاه در دایرکتوری موتور جستجوگر
- استراتژی URL
- بهینه سازی سورس
- پردازش کردن متن
- بهینه سازی لینک های ورودی
- استراتژی IP یا دامین

آنالیز تکنیک های بهینه سازی موتور جستجو

این پژوهش استراتژی بهینه سازی شرکت های گوگل و مایکروسافت را بررسی و توضیح داده است و موردی که جهت مطالعه و آنالیز قرار داده است، استراتژی های مربوط به شرکت مایکروسافت است. مهمترین معیارها برای بهینه سازی موتور جستجو از نظر محقق عبارتند از:

- اضافه کردن Allow یا DisAllow به نقشه دستگاه
- اضافه کردن URL به نقشه دستگاه
- اضافه کردن نقشه سایت به robots.txt
- استفاده از ابزارهای آنالیز جهت crawl

بهینه سازی موتور جستجو جهت افزایش بهبود رویت وب سایت

در این تحقیق محققین عواملی که در افزایش رتبه سایت تاثیر دارند را به 3 سه بخش تقسیم بندی می کنند:

- شرکت های دارنده موتور جستجو: جستجوکنندگان در صفحات موتورهای جستجو یک پرس و جو را اجرا می کنند سپس نتیجه آن در صفحه نتایج موتور جستجو نمایش داده می شود.
- وب مسترها و شرکت های بهینه سازی موتور جستجو: شرکت هایی که با استفاده از یکسری ابزار و مهارت ها تلاش می کنند رتبه وب سایت ها را در صفحه نتایج موتور جستجو افزایش دهند
- جستجو کنندگان در موتورهای جستجو: افرادی هستند که در موتورهای جستجو دنبال یافتن اطلاعات مورد نظرشان هستند.

محققین معتقد هستند افرادی که فعالیت بهینه سازی را انجام می دهند می بایست از آخرین تغییرات الگوریتم ها ی موتورهای جستجو مطلع باشند و عواملی که باعث می شود ترافیک سایت افزایش پیدا نماید را بخوبی بشناسند و همچنین از کلماتی که کاربران نهایی دنبال آن می گردند نیز آگاهی داشته باشند و عبارتی دیگر می بایست اطلاعات خود را در سه دسته بندی بالا همیشه بروز نگه دارند.

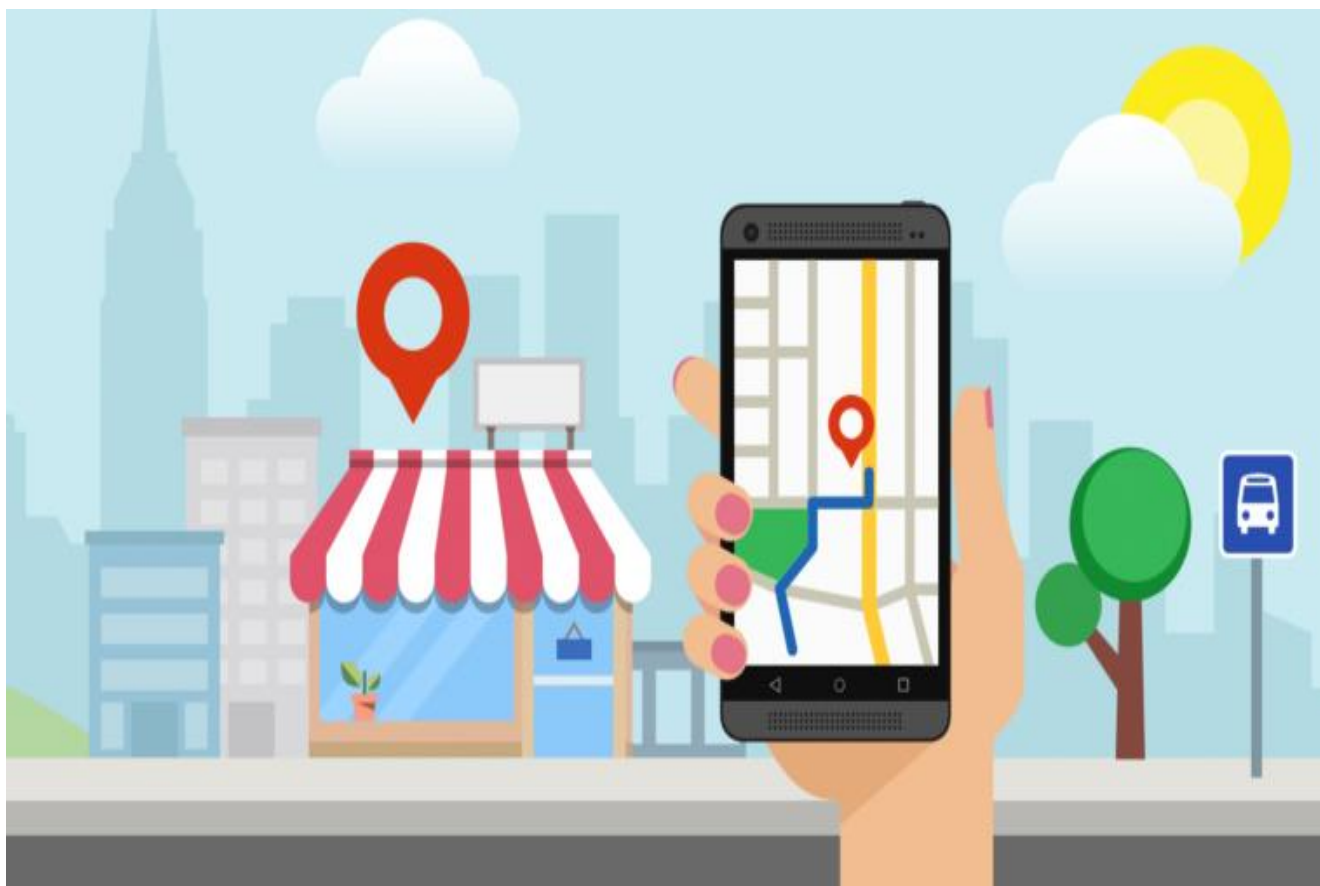
بهینه سازی موتور جستجو: مفاهیم آنالیز فنی

در این تحقیق بر اساس یکسری مقالات، مطالعات و راهنمایی ها برخی از مفاهیم بهینه سازی موتور جستجو آورده شده است. محقق معتقد هستند بهینه سازی موتور جستجو فرآیندی است که هیچ وقت تمام نمی شود و تکنیک های زیادی برای بهتر کردن رتبه سایت وجود دارد ولی پیاده سازی همه آنها مشکل است. :

تکنیک های بهبود رتبه دستگاه در نتایج جستجو ها

- افزایش زمان uptime
- دسترسی پذیری
- موقعیت جغرافیایی
- کلمات کلیدی هنگام رجستر کردن سایت
- استفاده از Meta Tag
- استفاده صحیح از عنوان
- بهینه سازی محتوا

فصل ششم الگوریتم های سئو برای داشتن بهترین نتایج



شکل 3-5: محلی سازی نتایج جستجوها

6-1 سئوی محلی یا Local Seo

در اینترنت اشیاء معمولاً اکثر نتایجی که دنبال آن هستیم در نزدیکی خود ما قرار دارند فرض بکنید شما میخواهید نزدیک ترین ایستگاه کرایه دوچرخه را که دوچرخه های آن خالی هستند را در یک موتور جستجوگر سرچ بکنید و نتایج آن را ببینید. برای اینکار باید در بین همه ایستگاه های کرایه دوچرخه ، نزدیک ترین به موقعیت مکانی شما یافت شود و بررسی شود که در این نزدیک ترین ها آیا دوچرخه هایی برای کرایه وجود دارد یا خیر.

این مورد قبلاً توسط گوگل نیز راه اندازی شده است و گوگل آن را با اسم الگوریتم کبوتر منتشر کرده بود.



شکل 4-5: نمونه الگوریتم نمایش نزدیکترین نتایج

2-6 الگوریتم کبوتر

این الگوریتم توسط گوگل ارائه شده است و برای جستجو کردن در یک محدوده خاص میباشد و بسیار کاربردی و مورد نیاز نیز میباشد. در اینترنت اشیاء هم معمولا کاربران یا کلاینت دنبال دریافت سرویس از نزدیکی خود هستند بخاطر همین کسی که در موتور جستجوگر دنبال دوچرخه کرایه ای میباشد، اگر در نتایج جستجو ها برای آن شخص اطلاعاتی درباره دوچرخه های کرایه ای موجود در کشوری دیگر نمایش داده شود پس موتور جستجوی ما نتوانسته دقیقا نیاز کاربر را رفع بکند.

مثال: فرض کنید در منطقه ونک در شهر تهران قرار دارد، اگر بخواهید بصورت آنلاین غذا سفارش بدهید با توجه به اینکه لیست سایت های فروش آنلاین غذا را ندارید بخاطر همین مجبور به جستجو کردن در گوگل خواهید بود، و عبارت "خرید غذا" یا "رستوران آنلاین" راسرچ میکنید و گوگل هم یکسری سایت را به عنوان نتایج برای شما نمایش میدهد اما اگر شما در نتایج لیست رستوران هایی نمایش داده شود که در شهر های تبریز، اصفهان و مشهد هستند پس به این نتیجه میرسیم که گوگل نتوانسته نیاز کاربر را رفع بکند.

برای رفع این مشکل گوگل یک الگوریتم مخصوص دارد که بر اساس موقعیت مکانی کاربر و نوع کلمه و عبارتی که جستجو کرده است مناسب ترین سایت ها را برای کاربر نمایش مید

نتیجه گیری

در نهایت پس از انجام مراحل بالا و تجزیه و تحلیل نکات گفته شده در مورد ساخت موتور جستجوگر در اینترنت اشیاء و نحوه سئو و بهینه سازی دستگاه ها برای این موتور جستجوگر باید به نتایج زیر رسیده باشیم:

1. توضیح کامل درباره اینترنت اشیاء را مطالعه نموده ایم
2. شناخت کامل درباره موتور های جستجو را پیدا کرده ایم
3. تاثیر اینترنت اشیاء را در زندگی خود احساس کرده ایم
4. احساس نیاز جامعه و محیط به راه اندازی موتور جستجوگر در اینترنت اشیاء
5. خزنده ها و نحوه ایندکس کردن سایت ها در موتور های جستجوگر
6. پیش نیاز های راه اندازی یک موتور جستجوگر در اینترنت اشیاء را شناخته ایم.
7. Block Chain و تاثیر آن بر روی اینترنت اشیاء
8. الگو گرفتن از Block Chain در رجستر کردن دستگاه ها در اینترنت اشیاء
9. موارد کاربرد بلاک چین اینترنت اشیا
10. توضیح کلی درباره پایگاه داده های سیار
11. کاربرد پایگاه داده های سیار در محلی سازی نتایج موتور های جستجو
12. چرا پایگاه داده های سیار
13. نیازمندیهای پایگاه داده های سیار
14. معماری سیستم mobile database
15. مراحل محلی سازی موتور جستجوی IoT
16. ابر ها در اینترنت اشیاء
17. مشخصات اینترنت اشیاء بر مبنای محاسبات ابری
18. انواع موتورهای جستجو
19. بررسی یک موتور جستجوی پیمایشی
20. نکات مهم در ساخت موتور جستجو
21. عامل موفقیت موتورهای جستجو
22. آمار استفاده از موتور های جستجو در کشور های مختلف
23. هوش مصنوعی در پیاده سازی موتور جستجوگر IoT
24. ویژگی های خزنده خوب

25. سنو یا بهینه سازی سایت چیست

26. سنوی محلی یا لوکال سنو

آینده اینترنت اشیاء

موتورهای جستجوگر رقابت سختی را خواهند داشت و گوگل هم اکنون هزینه و سرمایه گزاری زیادی را برای در دست گرفتن حوزه اینترنت اشیاء انجام داده است و با توجه به تعداد بسیار زیاد دستگاه ها و اشیاء در بین کاربران هم برای دریافت ورودی های بیشتر از طریق موتورهای جستجو رقابت بسیار سختی وجود خواهد داشت.

شرکت ها تلاش می کنند تا خدمات خود را به صورت مبتنی بر صدا در اینترنت اشیاء به مشتریان ارائه دهند. بر اساس لیست Fortune 500 پیچیدگی، گستردگی و کیفیت خدمات مبتنی بر صدا در سال 2018 رشد کرده و با توجه به خدمات موجود، حدوداً دو برابر خواهد شد. خدمات مالی و سایر صناعی که برای هر کاری بیش از یک کار ساده خواستار مجوز هستند عقب خواهند ماند. توسعه‌ی راهکارهای لبه در اینترنت اشیاء افزایش خواهد یافت.

امروزه فرآیندهای کسب و کار اینترنت اشیاء عمدتاً در مراکز داده‌ی محلی یا در ابر اتفاق می افتند. در سال 2018، شاهد حرکتی قابل توجه میان شرکت‌های توسعه‌دهنده‌ی این فرآیندها خواهیم بود، که به تحلیل داده به صورت محلی و در نزدیکی دستگاه‌های متصلی که این فرآیندها را به وجود می آورند نیاز خواهند داشت. دستگاه‌های لبه‌ی اینترنت اشیاء می توانند بر اساس داده‌ای که تولید می کنند، به صورت محلی عمل کرده و همچنین از مزایای ابر برای داشتن امنیت، مقیاس پذیری، پیکربندی، گسترش و مدیریت به خوبی بهره‌مند شوند.

سازندگان پلتفرم‌های صنعتی اینترنت اشیاء از کسب و کار IaaS خارج خواهند شد.

در طول ۱۲ ماه گذشته، تمام پلتفرم‌های صنعتی بزرگ اینترنت اشیاء برای ایجاد برخی از قابلیت‌های مختص صنعت یا اینترنت اشیاء از طریق تأمین‌کننده‌های خدمات ابری شامل آمازون، IBM و مایکروسافت اقدام کرده‌اند. همچنانکه این تأمین‌کننده‌ها برد جغرافیایی خود را گسترش می‌دهند، با قوانین دقیق نظارتی انطباق بیشتری می‌یابند و قابلیت‌های خود را در حوزه‌ی اینترنت اشیاء تقویت می‌کنند این روند نیز شتاب خواهد گرفت.

دستورالعمل‌های جدید اروپایی به سمت تجاری سازی داده‌های اینترنت اشیاء متمایل خواهند بود.

به گفته ی 45 درصد از تحلیلگران داده ها در شرکت های آمریکایی، آنها تا کنون داده های خود را تجاری سازی کرده اند؛ در حالیکه در اروپا، فقط 35 درصد از این شرکت ها در فرانسه و 38 درصد از آنها در آلمان این کار را انجام داده اند. با تشخیص این عقب ماندگی، کمیسیون اروپایی دستورالعمل هایی برای تشویق استفاده از تکنولوژی پیشرفته و رونق اقتصاد داده منتشر خواهد کرد.

بازاریاب ها از فرصت های موجود در اینترنت اشیاء آگاه می شوند، در حالیکه تکنولوژی های پوشیدنی در موقعیت خود باقی می مانند.

پذیرش فزاینده و فراگیر عامل های هوشمند مانند آمازون الکسا یا دستیار گوگل در بیشتر دستگاه ها، راه های جدیدی برای تعامل با مشتری به بازاریاب ها نشان خواهد داد. با ایجاد یک تجربه خوب از یک نام تجاری، بازاریاب ها استراتژی خود را برای داشتن رابط های جدید با استفاده از اسپیکرهای خانه هوشمند یا ساعت های هوشمند توسعه خواهند داد. با وجود فروش فزاینده بیش از 12 میلیون ساعت هوشمند، تنها در آمریکا تا پایان سال 2018، Forrester انتظار دارد میزان استفاده از تکنولوژی های پوشیدنی در موقعیت خود باقی بماند.

فهرست منابع

- [1]<http://nlp.stanford.edu/IR-book/html/htmledition/the-web-graph-1.html>
- [2]<http://www-personal.ksu.edu/~edderly/Google.pdf>
- [3]J. A. Aslam, and M. Montague, "Models for Metasearch". In Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM Press, New Orleans, USA, September 2001, 2001, pp. 276-284.
- [4]B. T. Bartell, "Optimizing Ranking Functions: A Connectionist Approach to Adaptive Information Retrieval", PhD Thesis, University of California, San Diego, USA, 1994.
- [5]McBryan, O.A. —GENVL and WWW: Tools for Taming the Web, In Proceedings of the First International World Wide Web Conference, pp. 79-90, 1994.
- [6]Gulli, A., Signorini, A., —The Indexable web is more than 11.5 billion pages, In Proceedings of the 14th international conference on World Wide Web, pp. 902-903, ACM Press, 2005.
- [7]<http://www.internetnews.com/stats/article.php/1363881>
- [8]http://www.pewinternet.org/pdfs/PIP_Searchengine_users.pdf
- [9]Baeza-Yates, R., & Ribeiro-Neto, B. (1999). Modern Information Retrieval. ACM Press/ Addison-Wesley.
- [10]Salton, G., & Buckley, C. (1988). Term weighting approaches in automatic text retrieval. Information Processing and Management, 24(5), 513–523.

- [11] Robertson, S. E., Walker, S., HancockBeaulieu, M. M., Gatford, M., & Payne, A. (1995). Okapi at TREC-4. In NIST Special Publication. The Fourth Text Retrieval Conference (TREC-4) (pp. 73—96).
- [12] Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). The pagerank citation ranking: Bringing order to the web. Technical report, Computer Science Department, Stanford University.
- [13] Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5), 604– 632.
- [14] Zareh Bidoki, A. M., & Yazdani, N. , DistanceRank: An intelligent ranking algorithm for web pages, *Information Processing and Management* (2007), doi:10.1016/j.ipm.2007.06.004.
- [15] Najork, M., Zaragoza, H., & Taylor, M. J. (2007). Hits on the web: how does it compare? In *Proceedings of SIGIR'07* (pp. 471-478).
- [16] Yager, R.R. (1988). On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Transactions on Systems, Man and Cybernetics*, 18(1).
- [17] Filev D. and Yager R. R., —On the issue of obtaining OWA operator weights, *Journal of Fuzzy Sets and Systems*, Vol. 94, No. 2, pp. 157-169, 1998.
- [18] Cho, J., Roy, S., & E. Adams, R. (2005). Page Quality: In Search of an Unbiased Web Ranking. In *Proceedings of ACM International Conference on Management of Data (SIGMOD)*.
- [19] Filev, D., and Yager, R. R., 1994. Learning OWA operator weights from data. *Proceedings. of the third IEEE Conference on Fuzzy Systems*, Vol.1, pp. 468-473.
- [20] LETOR, <http://research.microsoft.com/users/tyliu/LETOR/default.aspx>