همایش ملی الکترونیکی دستاوردهای نوین در علوم مهندسی و پایه
National e-Conference on Advances in Basic Sciences and Engineering
. . . . WWW.AEBSCONF.IR . . . .

امکان سنجی استفاده از تکنولوژی داده کاوی در بالابردن سطح امنیت داده ها در شرکتهای تابعه وزارت نفت

مهدی بازیار

آموزشکده فنی و حرفه ای سما ، دانشگاه آزاد اسلامی ، واحد کازرون ، کازرون ، ایران

چکیده

با توجه به وجود داده و اطلاعات بسیار زیاد در پایگاه داده شرکتهای تابعه نفت و اینکه تمامی داده مهم و حائز اهمیت برای شرکت می باشند که برخی از آنها داراس سطوح امنیتی بالا می باشند، لازم است با توجه به حملات سایبری چند وقت اخیر، بدنبال راه کارهایی باشیم تا بتوانیم بیش از پیش امنیت آنها را ایجاد سازیم و یکی از راه های ممکن شناسایی داده و اطلاعاتی می باشد که برای هکرها مفید بوده و بیشتر حملات معمولاً به این داده و اطلاعات رخ داده است. برای شناسایی و درجه بندی سطح علاقه مندی هکرها به انواع داده ها در این شرکتها، یکی از راه کارها استفاده از تکنولوژی داده کاوی می باشد. داده های موجود عموماً از نوع متنی و ارقام می باشند که در قسمتهایی همانند نامه ها، آمار مربوط به مشترکین و همکاران تجاری، اطلاعات متفاوت واحد منابع مالی و انسانی، لیست تجهیزات و منابع مورد نیاز اولیه و محل فیزیکی قرار گرفتن آنها و ... می باشد. در این مقاله به منظور بالابردن ضریب امنیتی داده ها ، مطالعه جامعی بر روی تکنیک جدید داده کاوی انجام شده و کاربرد و مزایا آنرا در بالابردن ضریب امنیت داده و اطلاعات در شرکتهای تابعه نفت معرفی شده است. در نهایت با توجه به تحقیقات بعمل آمده، چنین نتیجه گیری شده است که برای امن نمودن داده و اطلاعات در پایگاه داده به نحوی که ابتدا داده و اطلاعاتی که مورد نیاز هکرها می باشد را شناسایی و سپس با تدابیر ویژه ضریب امنیت آنها را بالا از روش داده کاوی استفاده شده است.

واژه‌های کلیدی

شرکتهای تابعه وزارت نفت ، امنیت داده و اطلاعات، داده کاوی

مقدمه

در شرکتهای تابعه نفت انواع داده های زیادی وجود دارد که معمولاً دارای اهمیت بالا می باشند و ضرورت برقراری امنیت بالا در سرور و پایگاه داده ها، بسیار مهم می باشد. داده های موجود که معمولاً از نوع متن می باشند حاوی اطلاعات مشترکین جزء و عمده و همچنین شرکای تجاری، نامه های با رده های مختلف سطح امنیت، اطلاعات پرسنل و تجهیزات، اطلاعات منابع مالی و انسانی، غیره می باشدکه این داده و اطلاعات در محیط های اینترانت داخلی شرکت و همچنین اینترنت در حال ذخیره، بازیابی، استفاده و جابجایی می باشند. با توجه به اینکه بسیاری از این اطلاعات از لحاظ شرکت دارای اهمیت بالایی می باشند، بایستی روشهایی را پیدا و استفاده کنیم تا حداکثر امنیت لازم در زمان جابجایی اطلاعات در محیط های با سیم و بدون سیم را برقرار سازیم با توجه به این مساله حساس و مهم، بایستی از تکنیک داده کاوی استفاده شود که دارای خصوصیات زیر می باشد:

1.فرایندی است خودکار، برای استخراج الگوهایی که دانش را بازنمایی می کنند،که این دانش ها بصورت ضمنی در پایگاه داده های عظیم، انباره داده و دیگر مخازن بزرگ اطلاعات ذخیره شده اند. 2 .داده کاوی همزمان از چندین رشته علمی همانند تکنولوژی پایگاه داده، هوش مصنوعی، یادگیری ماشین، شبکه عصبی، آمار،شناسایی الگو، سیستم های مبتنی بر دانش، حصول دانش، بازیابی اطلاعات، محاسبات سرعت بالا و بازنمایی بصری داده ها استفاده می نماید. 3.تحت محدودیت های موثر محاسباتی قابل قبول، الگو و یا مدلها را در داده کشف می کند و کاربرد ویژه ای در حوزه های تصمیم گیری، پیش بینی، پیشگویی و تخمین زدن در داده های حجیم ولی بدون ارزش را دارد که اغلب به آن تحلیل داده ای ثانویه گویند.

هدف از این مقاله ارائه راهکارهایی جهت بالابردن سطح امنیت داده و اطلاعات در این شرکتها می باشد و این نتیجه حاصل می شود.

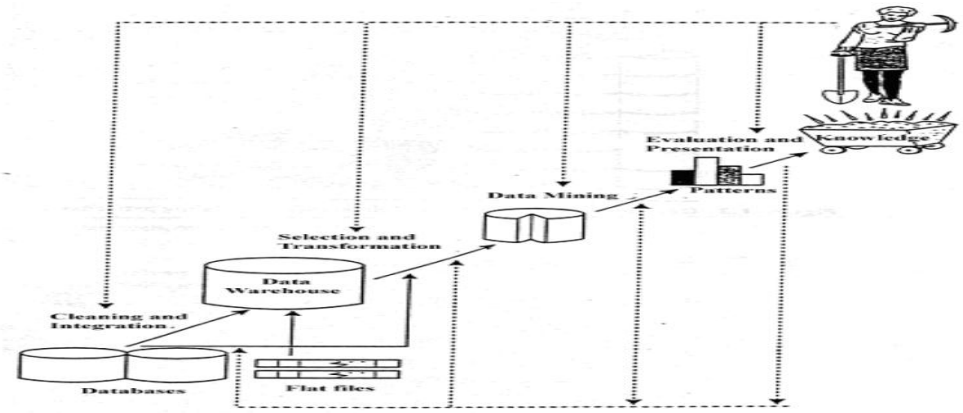
ساختار ادامه مقاله به قرار زیر است در بخش 2 به مروری بر تعاریف داده کاوی و کاربردهای آن ، در بخش 3 بررسی و ارزیابی و ارائه راهکارهای جدید، بخش 4 مربوط به نتیجه گیری می باشد.

مروری بر تعاریف داده کاوی و کاربردهای آن

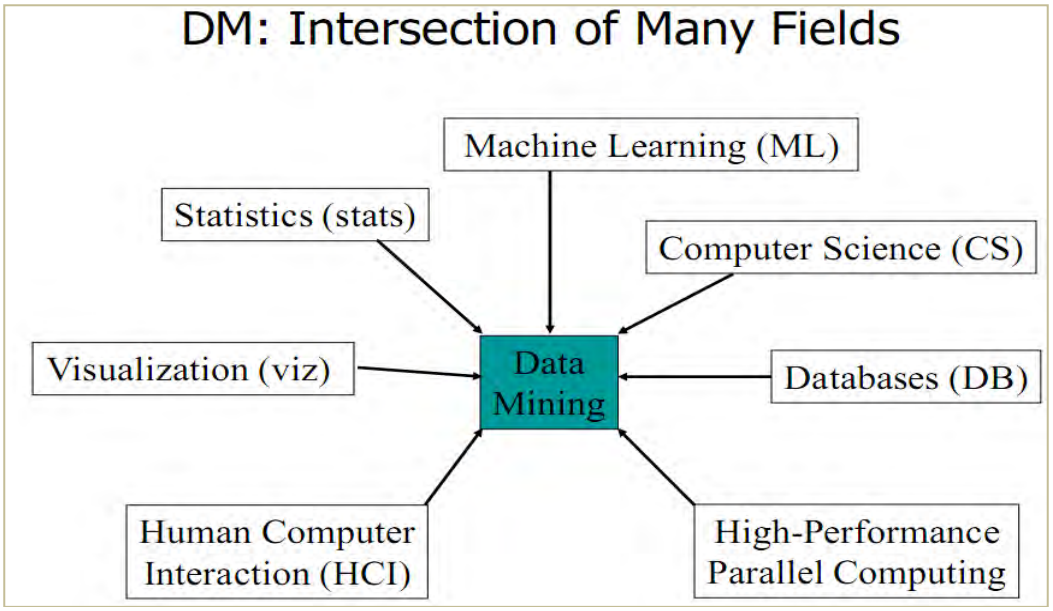
با توجه به حملات روزافزون هکرها که ممکن است کشور، شرکت و یا شخصی باشند که از طریق برنامه های مخرب متعدد به سیستم سرور و پایگاه داده اطلاعات شرکت وارد شده و بعضاً عملیاتی مخرب را انجام می دهند، لازم است با استفاده از تکنولوژی داده کاوی به سؤالات ذیل دست یافته و درنتیجه با استفاده از راهکارهای مناسب، امنیت داده و اطلاعات را افزایش دهیم:

آکادمی داده
dataacademy.ir

3) شناسایی مکان و یا اشخاص،کشور و یا شرکتهایی که تا به امروز به داده و اطلاعات دسترسی غیر مجاز داشته و حملاتی را انجام داده اند.
4) شناسایی درجه مخرب بودن حملات و مکانها ،اشخاص و یا کشورهایی که بیشترین حملات را به پایگاه داده شرکت داشته اند.
5) شناسایی آندسته داده و اطلاعات شرکت که بیشتر مورد توجه هکرها قرار داشته و بیشترین حملات را به آن اطلاعات داشته اند.
6) شناسایی راههای ممکن ورود برنامه های مخرب به سیستم های شرکت
7) شناسایی میزان ایمن بودن سرورهای شرکت در برابر حملات
8) در نهایت افزایش ضریب امنیت داده و اطلاعات پایگاه داده های شرکت با استفاده از نتایج حاصله از مراحل قبل و با استفاده از روش داده کاوی



شکل1 –نحوه کار داده کاوی



شکل2- منابع جمع آوری اطلاعات در داده کاوی

داده کاوی پل ارتباطی علم آمار، علم کامپیوتر، هوش مصنوعی، الگوشناسی، فراگیری ماشین و بازنمایی بصری داده می باشد. همچنین فرایندی پیچیده جهت شناسایی الگوها و مدلهای صحیح، جدید و بصورت بالقوه مفید، در حجم وسیعی از داده ها می باشد، به طریقی که این الگوها و مدلها برای انسانها قابل درک باشند. داده کاوی بصورت یک محصول قابل خریداری نمی باشد، بلکه یک رشته علمی و فرایندی است که بایستی بصورت یک پروژه ساده بررسی شود.

داده کاوی فرایندی تحلیلی است که برای کاوش داده ها(معمولاً حجم عظیمی از داده ها،در زمینه های کسب و کار و بازار)صورت گرفته و یافته ها با بکارگیری الگوها، دارای اعتبار لازم میشوند.هدف اصلی داده کاوی پیش بینی است.

می توان داده کاوی را به چنین شکل بیان کرد:

ر یک پایگاه داده، که با استفاده از پردازشهای معمول قابل دستیابی نیستند.

آکادمی داده
dataacademy.ir

1- داده کاوی عبارتست از فرایند استخراج اطلاعات معتبر، از پیش ناشناخته، قابل فهم و قابل اعتماد از پایگاه داده های بزرگ و استفاده از آن در تصمیم گیری در فعالیت های تجاری مهم

2- فرایند نیم خودکار تجزیه و تحلیل پایگاه داده های بزرگ به منظور یافتن الگوهای مفید

3- فرایند جستجو در یک پایگاه داده برای یافتن الگویی میان داده ها

4- تجزیه و تحلیل مجموعه داده های قابل مشاهده برای یافتن روابط مطمئن بین داده ها

5- استخراج دانش کلان، قابل استناد و جدید از پایگاه داده های بزرگ

با توجه به اینکه داده کاوی تئوریهای پایگاه داده ها، هوش مصنوعی، یادگیری ماشین و علم آمار را در هم می آمیزد و با توجه به استفاده از اطلاعاتی با حجم زیاد در حد گیگا و یا ترابایت، می توان از کاربردهای آن در شناسایی بزرگترین بازار هدف، انبار جامع داده ها، مراکز داده و سیستم های پشتیبانی تصمیم برای بدست آوردن تخصص هایی در صنایعی همانند شبکه های توزیع مویرگی، تولید، مخابرات، بیمه، زنجیره عرضه، کشف ضعفهای امنیتی سیستم برای تصمیم به منظور برطرف نمودن آنها و غیره اشاره نمود.

از فنون رایج بکار گرفته شده تحت عنوان داده کاوی می توان به موارد زیر اشاره نمود:

1- ابزارهای پرس و جو

2- فنون آماری

3- مصورسازی

4- پردازش تحلیلی پیوسته

5- یادگیری مبتنی بر مواد

6- درختان تصمیم

7- قوانین وابستگی

8- شبکه های عصبی

9- الگوریتم ژنتیک

اصولاً هرکجا که داده وجود دارد، داده کاوی نیز مطرح می گردد، از قبیل : امور پزشکی، امور تجاری و مالی، زیست پزشکی، تجزیه و تحلیل های مربوط به DNA، کشف ناهنجاریها و اسناد جعلی، ارتباطات از راه دور، ورزش و سرگرمی، کتابداری و اطلاع رسانی. همچنین در تمامی شرکتهایی که مشتریان در کانون توجه قرار می گیرند.

با توجه به اینکه داده کاوی پیش بینی وضع آینده بازار، گرایش مشتریان و شناخت سلیقه های عمومی آنها را برای شرکتها ممکن می سازد می توان به موارد کاربرد آن اشاره نمود:

1- بازار هدف 2- پیدا کردن الگوی خرید مشتری 3- برنامه ریزی برای معرفی محصول جدید 4 -Customer Profiling 5- دسته بندی مشتریان بر اساس نوع خرید 6- آنالیز نیازهای مشتریان 7- تشخیص محصولات مناسب برای دسته های مختلف مشتریان 8- تشخیص فاکتورهایی برای جذب مشتریان جدید 9- تعیین الگوهای خرید مشتریان 10- تجزیه و تحلیل سبد خرید بازار 11- پیشگویی میزان خرید مشتریان از طریق پست(فروش الکترونیکی) 12-پیش بینی الگوهای کلاهبرداری از طریق کارتهای اعتباری و شناسایی جرایم مالی 13-شناسایی داده و اطلاعاتی که دارای ضریب پایین به منظور جلوگیری از نفوذ افراد و برنامه های مخرب می باشند 14-شناسایی آندسته از داده و اطلاعات شرکت که برای هکرها برای نفوذ به درون آنها و استفاده غیر مجاز از آنها بسیار مهم می باشد 15- تشخیص مشتریان ثابت و دسته بندی و خوشه بندی مشتریان با توجه به رفتار مشابه آنها در زمینه بانکداری و بازپرداخت وام 16- تعیین میزان استفاده از کارتهای اعتباری بر اساس گروههای اجتماعی 17- تحلیل اعتبار مشتریان 18- شناسایی فاکتورهای اصلی در ریسک بازپرداخت وام 19- تحلیل پاسخگویی مشتریان به ارائه خدمات جدید بانکی 20- پیشگویی میزان خرید بیمه نامه های جدید توسط مشتریان 21- تحلیل ریسک و برآورد حق بیمه مشتریان بر اساس میزان ریسک هر مشتری 22- پیش بینی میزان خسارت بر اساس گروههای مشترین 23- مدیریت ارتباط با بیمه گذاران و تدوین استراتژی بر اساس مشتریان هدف 24- تعیین عوامل وفاداری و یا رویگردانی مشتریان 25- شناخت نیازها و الگوهای خرید سرویسهای بیمه ای توسط مشتریان 26- شناخت تلفات بیمه ای 27- تعیین نوع رفتار با بیماران و تعیین روش درمان بیماریها 28- پیشگویی میزان موفقیت اعمال جراحی و تعیین میزان موفقیت روشهای درمانی در برخورد با بیماریهای سخت 29- بررسی میزان تاثیر داروبر بیماری و اثرات جانبی آن

داده کاوی حاصل سیر تکاملی طبیعی تکنولوژی اطلاعات در صنعت پایگاه داده نظیر : جمع آوری داده ها و ایجاد پایگاه داده، مدیریت داده و تحلیل و فهم داده ها می باشد.

احل داده کاوی

آکادمی داده
dataacademy.ir

مراحل کشف دانش با استفاده از تکنیک داده کاوی به شرح زیر می باشد:

1- پاکسازی داده(ازبین بردن نویز و ناسازگاری داده) 2- یکپارچه سازی داده 3- انتخاب داده های مورد نظر جهت آنالیز نمودن 4- تبدیل نمودن داده ها(خلاصه سازی و همسان سازی به فرم مناسب جهت داده کاوی) 5- داده کاوی با استفاده از روشهای هوشمند 6- ارزیابی الگو بوسیله معیارهای اندازه گیری 7- ارائه دانش

داده کاوی دارای دو طبقه بندی توصیفی و پیش گویانه می باشد.

انواع مدلهای پیش بینی:

Time Series ،Regression ،Classification

هوانگ و همکاران (2002) گزارش نمودند، با استفاده از تکنیک داده کاوی می توان جمعیت شناختی درستی را انجام داد و نسبت به این اطلاعات، محصولات متفاوت و مورد نیاز هر منطقه را طراحی و آماده نمود.

سیمیونیدیس و همکاران(2006)گزارش نمودند، با استفاده از تکنیک داده کاوی می توان بهترین مدیریت زنجیره تامین (SCM) را جهت شناسایی انواع محیطها با پتانسیلهای متفاوت و بالا برای انجام بهترین مزایده و یا مناقصه انجام داد.

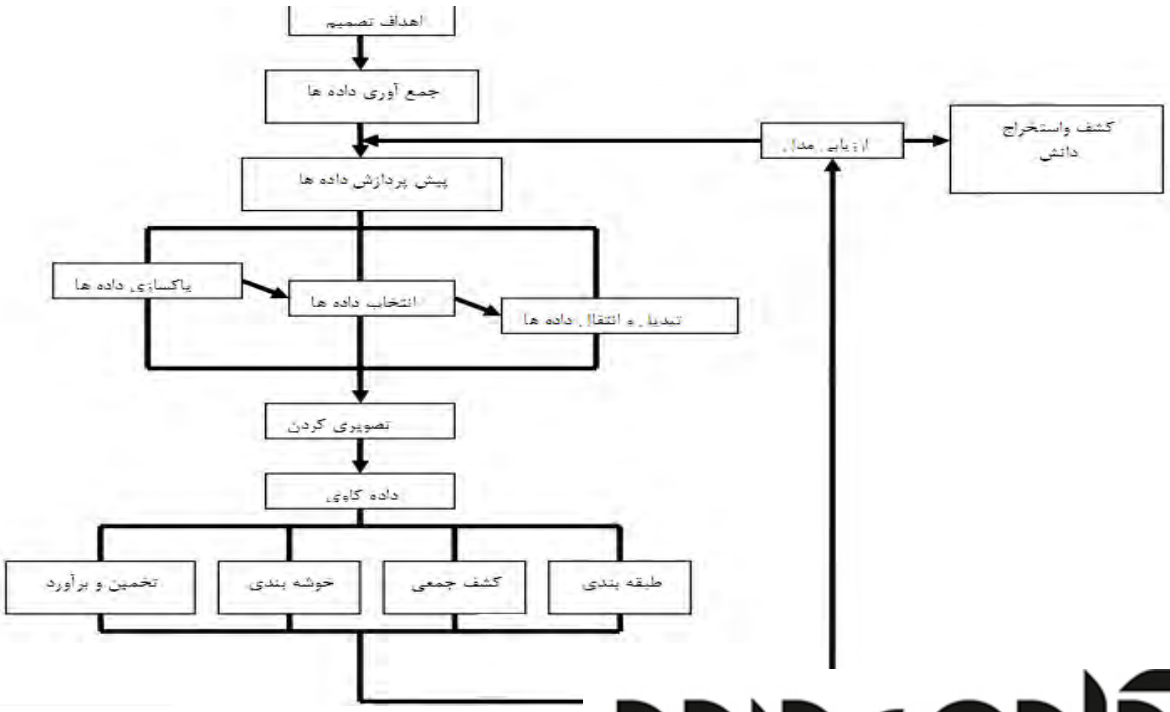
جارگ و همکاران(2011) گزارش نمودند، استفاده از تکنیک داده کاوی جهت شناسایی حملات در شبکه های جهانی همانند اینترنت بسیار کاربردی و مهم می باشد. آنها با جمع آوری اطلاعات بسیار زیاد و استفاده از این تکنولوژی دریافتند که حمله ای به نام Zombies isa در اینترنت و برروی بسیاری از وب سایتها همانند یاهو، سی ان ان و آمازون انجام گرفته شده است.

نگوین و همکاران(2008) گزارش نمودند، در چند سال گذشته حملات به شبکه بسیار زیاد شده و بهمین دلیل استفاده از سیستم تشخیص نفوذ(IDS) به طور فزاینده ای به یک جزء حیاتی برای حفظ شبکه تبدیل شده است. با توجه به حجم زیادی از اطلاعات امنیتی و همچنین خواص پیچیده و پویا از رفتارهای نفوذ، بهینه سازی هملکرد IDS بسیار مشکل شده است و برای بهینه سازی آن استفاده از تکنیک داده کاوی و الگوریتم های طبقه بندی برروی مجموعه ای از داده ها، می تواند کاربردی باشد.

بلویدورن و همکاران (2000) گزارش دادند، با توجه به حملات بسیار زیاد اتفادق افتاده در شبکه های کامپیوتری، می بایست با استفاده از تکنیک داده کاوی و استفاده از دو گروه : متخصصان امنیت شبکه با آشنایی کمی از تکنیک داده کاوی و کارشناسان داده کاوی با آشنایی کمی در سیستم تشخیص نفوذ شبکه، انواع حملات را شناسایی و راهکارهای مناسب جهت جلوگیری از آنها را اتخاذ نمود.

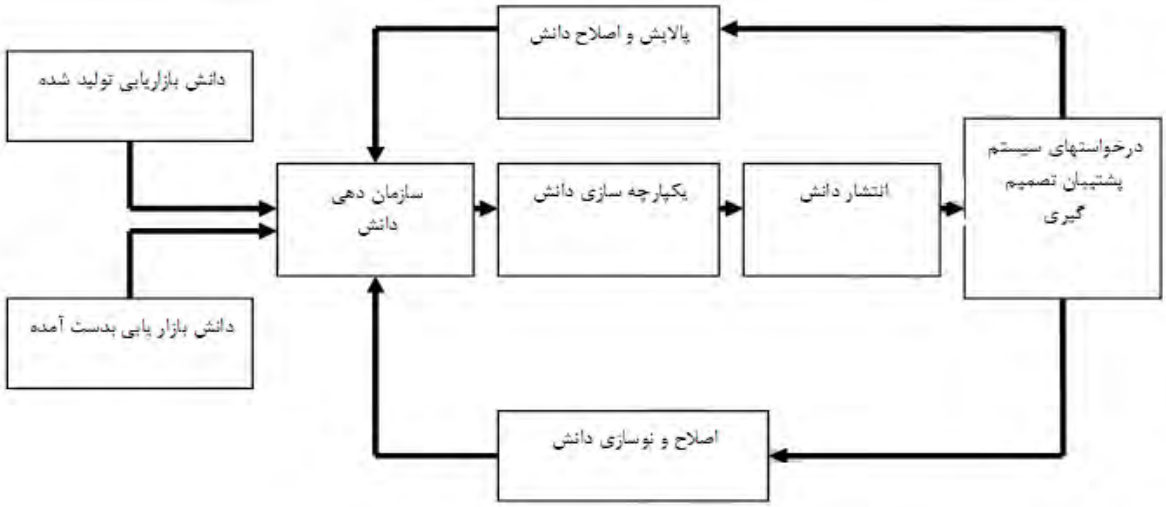
نایک(2008) گزارش داد، برای کسب و کار مناسب بایستی بینش مناسب از وضعیت فعلی داشته باشیمتا بتوانیم فرصت های جدیدی جهت ارائه خدمات بهتر به مشتریان را شناسایی کنیم. برای اینکار بایستی به حقایقی دست یابیم که ناشناخته هستندو برای بدست آوردن آنها لازم است که در انبوه زیادی از داده و اطلاعات بدنبال نتایج دلخواه بگردیم. استفاده از تکنیک داده کاوی که برای پالایش اطلاعات نیاز به استفاده از مجموعه اطلاعات پایگاه داده، وب سایتها، مشتریان، تکنولوژی های روز، سلایق و علایق بازار فروش ، کشف وب سرویس های ایده آل، دسته بندی شباهت ها و داده های توزیع شده می باشد می تواند بسیار مفید و صرفه جویی قابل توجهی در هزینه کسب و کار را فراهم آورد.

بررسی و ارزیابی و ارائه راهکارهای جدید



د استخراج دانش

آکادمی داده
dataacademy.ir



شکل5–سازماندهی،انتشار و اصلاح دانش

نتیجه گیری

با توجه به مقالات بررسی شده، چنین نتیجه حاصل می شود که تکنولوژی داده کاوی با بررسی حجم بسیاری از داده ها ، می تواند آمار و اطلاعاتی را به ما بدهد که این اطلاعات در مرحله اول قابل کشف توسط انسان نبوده و بسیار حائز اهمیت می باشند همانند اینکه درصد نفوذ و حمله به سرورها، از چه کشورهایی می باشد. چند درصد حملات برای جاسوسی بوده و یا بدون تاثیر بوده اند. معمولاً در چه بازه های زمانی حملات اتفاق افتاده است. نحوه نفوذ و حمله به داده ها به چه طریقی بوده و در نهایت اینکه چه نوع داده هایی بیشتر مورد توجه برای حمله بوده است.

منابع مورد استفاده

Hwang , S. , Lim , E. 2002. A Data Mining Approach to New Library Book Recommendations. ICADL, LNCS 2555, pp. 229-240, 2002 © Springer-Verlag Berlin Heidelberg 2002

Symeonidis , A.L. , Nikolaidou , V. , Mitkas ,P. A. 2006. Exploiting Data Mining techniques for improving the efficiency of a Supply Chain Management agent . WI-IATW '06 Proceedings of the 2006 IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology Pages 23-26 .IEEE Computer Society Washington, DC, USA ©2006

Garg , K. , Chawla ,R. 2011. DETECTION OF DDOS ATTACKS USING DATA MINING. International Journal of Computing and Business Research (IJCBR) ISSN (Online) : 2229-6166 Volume 2 Issue 1

Bloedorn , E. , Christiansen , A. D. , Hill , W. , Skorupka , C. , Talbot , L. M. , Tivel , J. 2000.

Nayak , R. 2008. Data Mining in Web ServicesDiscovery and Monitoring. International Journal of Web Services Research , Vol.X, No.X, 200X

Nguyen , H. A. , Choi , D. 2008. Application of Data Mining to Network Intrusion Detection: Classifier Selection Model . APNOMS 2008, LNCS 5297, pp. 399–408, 2008.© Springer-Verlag Berlin Heidelberg 2008

آکادمی داده
dataacademy.ir