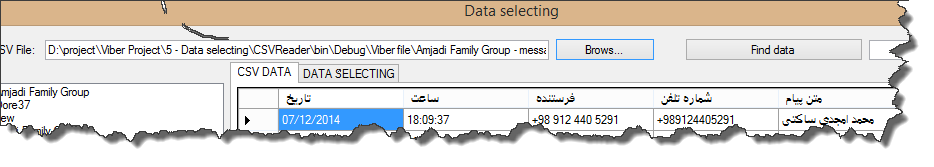


|  |
| --- |
| **بسمه تعالی** |
|  |
| **عنوان پروژه:** |
| آنالیز انتشار داده ها در شبکه های اجتماعی |
|  |
| **مبحث پروژه:** |
| رایانه |
|  |
| **پژوهشگر:** |
| امیرحسین زارعی |
|  |
| **استاتید راهنما:** |
| جناب آقای محمد مهدی رستگار و محمد قاسم نژاد |
|  |
| C:\Users\AMIRHO~1\AppData\Local\Temp\SNAGHTML2a4ed442.PNGسال تحصیلی 94-95 |



فهرست مطالب

[**کلید واژگان** 4](#_Toc460868183)

[**چکیده مطلب** 5](#_Toc460868184)

[**مقدمه** 6](#_Toc460868185)

[**بیان مسئله** 8](#_Toc460868186)

[**پیشینه تحقیق و قوانین علمی و قوانین نظری بکار رفته** 10](#_Toc460868187)

[**شرح پروژه** 11](#_Toc460868188)

[**شرح قسمت وایبر پروژه(قسمت فرعی)** 12](#_Toc460868189)

[**بدست آوردن داده ها** 12](#_Toc460868190)

[**حریم خصوصی** 12](#_Toc460868191)

[**بخش پالایش محتوا پیام** 13](#_Toc460868192)

[**یافتن تشابه و نمایش صفات** 13](#_Toc460868193)

[**شرح پروژه قسمت تلگرام(قسمت اصلی)** 15](#_Toc460868194)

[**ربات خزنده(Robot crawling)** 15](#_Toc460868195)

[**بخش ظاهری:** 15](#_Toc460868196)

[**بخش باطنی:** 15](#_Toc460868197)

[**حریم خصوصی** 16](#_Toc460868198)

[**پالایش محتوای پیام** 16](#_Toc460868199)

[**یافتن تشابه و نمایش صفات** 16](#_Toc460868200)

[**ایجاد دیتابیس های موضوعات مختلف** 17](#_Toc460868201)

[**آنالیز و تحلیل ها** 20](#_Toc460868202)

[**شبیه سازی** 21](#_Toc460868203)

[**کنترل جامعه هدف** 23](#_Toc460868204)

[**کنترل تقریبی و ارسال پیام** 23](#_Toc460868205)

[**بات سواری** 23](#_Toc460868206)

[**سرور اجرایی ربات** 24](#_Toc460868207)

[**چشم انداز توسعه** 25](#_Toc460868208)

[**کاربرد طرح** 26](#_Toc460868209)

[**مزایای پروژه** 27](#_Toc460868210)

[**معایب پروژه** 27](#_Toc460868211)

[**منابع** 28](#_Toc460868212)

[**نمودار ها** 29](#_Toc460868213)

[نمودار مولفه های پروژه 29](#_Toc460868214)

[نمودار بدست آوردن دیتا در تلگرام 30](#_Toc460868215)

[نمونه ای از بخش نمودار های هر پیام در پروژه 31](#_Toc460868216)

[نمونه نمودار های کلی پیام ها در پروژه 32](#_Toc460868217)

[**ضمایم پروژه:** 33](#_Toc460868218)

[عکس هایی از پروژه: 33](#_Toc460868219)

[بخش نمایش دیتابیس های موضوعی: 33](#_Toc460868220)

[صفحه ابتدایی برنامه مربوط به تمام لاگ های رباط در هنگام خزش و جدیدترین منابع آپدیت دیتابیس موضوعی 34](#_Toc460868221)

[تصویری از محیط پروژه در ویژوال استدیو 2015 35](#_Toc460868222)

[مودلینگ یا طرح کلی پشت پرده ی دیتابیس ها 36](#_Toc460868223)

[تصاویری از برنامه خواندن اطلاعات از روی فایل پشتیبانی وایبر 37](#_Toc460868224)

[مقاله های مربوط پیشینه ی تحقیق: 38](#_Toc460868225)

[Bigdata چیست و چرا اهمیت دارد؟ 38](#_Toc460868226)

[ربات تلگرام چیست؟ 40](#_Toc460868227)

[تفاوت های ربات تلگرام با افراد 40](#_Toc460868228)

[Bots: An introduction for developers 41](#_Toc460868229)

[1. What can I do with bots? 41](#_Toc460868230)

[2. How do bots work? 42](#_Toc460868231)

[3. How do I create a bot? 42](#_Toc460868232)

[4. How are bots different from humans? 42](#_Toc460868233)

[5. Bot perks 43](#_Toc460868234)

[6. BotFather 50](#_Toc460868235)

[Telegram APIs 52](#_Toc460868236)

[Bot API 52](#_Toc460868237)

[Telegram API 52](#_Toc460868238)

[Delta encoding 56](#_Toc460868239)

[Levenshtein distance 56](#_Toc460868240)

[Edit distance 56](#_Toc460868241)

[RSS چیست ؟ 57](#_Toc460868242)

[مرتبط با رسم گراف در شبیه ساز 58](#_Toc460868243)

# **کلید واژگان**

* شبیه سازی(simulation)
* ربات خزنده(Robot crawling)
* آنالیز انتشار داده
* آنالیز متن
* Big data
* Weight word
* وزن کلمه
* **تلگرام(telegram)**
* وایبر(viber)
* داده ها، پیام ها، اطلاعات
* سرعت پخش پیام
* روند پخش پیام
* موضوعات و زمانبندی ها
* دیتابیس موضوعات
* کنترل جامعه ی هدف
* تحلیل نموداری
* شبیه سازی با گراف

# **چکیده مطلب**

با استفاده ی روز افزون از شبکه های اجتماعی به خصوص شبکه های اجتماعی فراگیر همچون تلگرام، وایبر و... با توجه به کثرت کاربران آن ها، روند پخش اطلاعات متناسب با ویژگی های ساختاری و جامعه ی هدف می باشد. بنابراین طراحی و تعیین ساختار مناسب پیام متناسب با هدف عامل اصلی کنترل روند پخش پیام می باشد. و همچنین هدف سازمان ها از حضور در فضای مجازی همین موضوع است. در این پروژه با تحلیل هوشمند ویژگی های ساختاری و زمانی پیام و روند پخش اطلاعات و ارتباط آن با پارامتر های مذکور، ساختار پیام متناسب با هدف آن بهینه شده و باعث میشود بتوان یک پیام را به صورت تقریبا کنترل شده ای در شبکه‌های اجتماعی پخش کرد همچنین با تحلیل های نموداری پیام ها و جوامع هدف و ارتباط مستقیم با آنها شرایط مناسب را برای کنترل یک جامعه ی هدف خاص فراهم می شود.

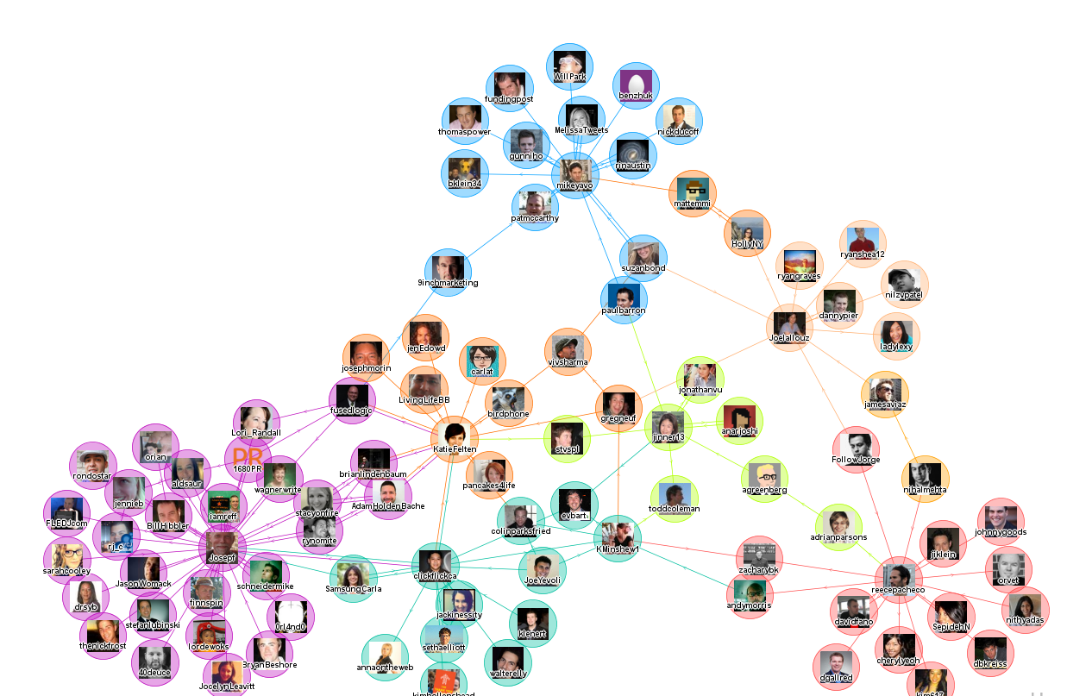
# **مقدمه**

آنالیز شبکه‌های اجتماعی مرتبط با [نظریه شبکه‌ها](http://fa.wikipedia.org/w/index.php?title=%D9%86%D8%B8%D8%B1%DB%8C%D9%87%D9%94_%D8%B4%D8%A8%DA%A9%D9%87&action=edit&redlink=1&preload=%D8%A7%D9%84%DA%AF%D9%88:%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%D9%85%D9%82%D8%A7%D9%84%D9%87/%D8%A7%D8%B3%D8%AA%D8%AE%D9%88%D8%A7%D9%86%E2%80%8C%D8%A8%D9%86%D8%AF%DB%8C&editintro=%D8%A7%D9%84%DA%AF%D9%88:%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%D9%85%D9%82%D8%A7%D9%84%D9%87/%D8%A7%D8%AF%DB%8C%D8%AA%E2%80%8C%D9%86%D9%88%D8%AA%DB%8C%D8%B3&summary=%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%DB%8C%DA%A9+%D9%85%D9%82%D8%A7%D9%84%D9%87+%D9%86%D9%88+%D8%A7%D8%B2+%D8%B7%D8%B1%DB%8C%D9%82+%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF%DA%AF%D8%B1&nosummary=&prefix=&minor=&create=%D8%AF%D8%B1%D8%B3%D8%AA+%DA%A9%D8%B1%D8%AF%D9%86+%D9%85%D9%82%D8%A7%D9%84%D9%87+%D8%AC%D8%AF%DB%8C%D8%AF&withJS=MediaWiki:Intro-Welcome-NewUsers.js) به عنوان یک تکنیک کلیدی در [جامعه‌شناسی](http://fa.wikipedia.org/wiki/%D8%AC%D8%A7%D9%85%D8%B9%D9%87%E2%80%8C%D8%B4%D9%86%D8%A7%D8%B3%DB%8C)، [انسان‌شناسی](http://fa.wikipedia.org/wiki/%D8%A7%D9%86%D8%B3%D8%A7%D9%86%E2%80%8C%D8%B4%D9%86%D8%A7%D8%B3%DB%8C)،  [جغرافیا](http://fa.wikipedia.org/wiki/%D8%AC%D8%BA%D8%B1%D8%A7%D9%81%DB%8C%D8%A7)، [روانشناسی اجتماعی](http://fa.wikipedia.org/wiki/%D8%B1%D9%88%D8%A7%D9%86%D8%B4%D9%86%D8%A7%D8%B3%DB%8C_%D8%A7%D8%AC%D8%AA%D9%85%D8%A7%D8%B9%DB%8C)، [جامعه‌شناسی زبان](http://fa.wikipedia.org/wiki/%D8%AC%D8%A7%D9%85%D8%B9%D9%87%E2%80%8C%D8%B4%D9%86%D8%A7%D8%B3%DB%8C_%D8%B2%D8%A8%D8%A7%D9%86)، [علوم ارتباطات](http://fa.wikipedia.org/w/index.php?title=%D8%B9%D9%84%D9%88%D9%85_%D8%A7%D8%B1%D8%AA%D8%A8%D8%A7%D8%B7%D8%A7%D8%AA&action=edit&redlink=1&preload=%D8%A7%D9%84%DA%AF%D9%88:%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%D9%85%D9%82%D8%A7%D9%84%D9%87/%D8%A7%D8%B3%D8%AA%D8%AE%D9%88%D8%A7%D9%86%E2%80%8C%D8%A8%D9%86%D8%AF%DB%8C&editintro=%D8%A7%D9%84%DA%AF%D9%88:%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%D9%85%D9%82%D8%A7%D9%84%D9%87/%D8%A7%D8%AF%DB%8C%D8%AA%E2%80%8C%D9%86%D9%88%D8%AA%DB%8C%D8%B3&summary=%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%DB%8C%DA%A9+%D9%85%D9%82%D8%A7%D9%84%D9%87+%D9%86%D9%88+%D8%A7%D8%B2+%D8%B7%D8%B1%DB%8C%D9%82+%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF%DA%AF%D8%B1&nosummary=&prefix=&minor=&create=%D8%AF%D8%B1%D8%B3%D8%AA+%DA%A9%D8%B1%D8%AF%D9%86+%D9%85%D9%82%D8%A7%D9%84%D9%87+%D8%AC%D8%AF%DB%8C%D8%AF&withJS=MediaWiki:Intro-Welcome-NewUsers.js)، [علوم اطلاعات](http://fa.wikipedia.org/wiki/%D8%B9%D9%84%D9%88%D9%85_%D8%A7%D8%B7%D9%84%D8%A7%D8%B9%D8%A7%D8%AA)، [مطالعات سازمانی](http://fa.wikipedia.org/w/index.php?title=%D9%85%D8%B7%D8%A7%D9%84%D8%B9%D8%A7%D8%AA_%D8%B3%D8%A7%D8%B2%D9%85%D8%A7%D9%86%DB%8C&action=edit&redlink=1&preload=%D8%A7%D9%84%DA%AF%D9%88:%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%D9%85%D9%82%D8%A7%D9%84%D9%87/%D8%A7%D8%B3%D8%AA%D8%AE%D9%88%D8%A7%D9%86%E2%80%8C%D8%A8%D9%86%D8%AF%DB%8C&editintro=%D8%A7%D9%84%DA%AF%D9%88:%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%D9%85%D9%82%D8%A7%D9%84%D9%87/%D8%A7%D8%AF%DB%8C%D8%AA%E2%80%8C%D9%86%D9%88%D8%AA%DB%8C%D8%B3&summary=%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%DB%8C%DA%A9+%D9%85%D9%82%D8%A7%D9%84%D9%87+%D9%86%D9%88+%D8%A7%D8%B2+%D8%B7%D8%B1%DB%8C%D9%82+%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF%DA%AF%D8%B1&nosummary=&prefix=&minor=&create=%D8%AF%D8%B1%D8%B3%D8%AA+%DA%A9%D8%B1%D8%AF%D9%86+%D9%85%D9%82%D8%A7%D9%84%D9%87+%D8%AC%D8%AF%DB%8C%D8%AF&withJS=MediaWiki:Intro-Welcome-NewUsers.js)، [اقتصاد](http://fa.wikipedia.org/wiki/%D8%B9%D9%84%D9%85_%D8%A7%D9%82%D8%AA%D8%B5%D8%A7%D8%AF) و [زیست‌شناسی](http://fa.wikipedia.org/wiki/%D8%B2%DB%8C%D8%B3%D8%AA%E2%80%8C%D8%B4%D9%86%D8%A7%D8%B3%DB%8C) مدرن همانند یک موضوع محبوب در زمینه تفکر ومطالعه پدیدار شده‌است.

بالغ بر یک قرن است که مردم، شبکه اجتماعی مجازی را برای اشاره‌های ضمنی به مجموعه روابط پیچیده میان افراد درسیستم‌های اجتماعی در تمامی مقیاس‌ها از روابط بین فردی گرفته تا بین‌المللی مورد استفاده قرار می‌دهند. در سال ۱۹۴۵، *J. A. Barnes*برای نخستین بار از اصطلاح قاعده‌مند برای مشخص کردن الگوهایی از رشته‌ها استفاده کرد که مفاهیم را مشخص می‌کنند و به صورت رایج توسط عموم و دانشمندان [علوم اجتماعی](http://fa.wikipedia.org/wiki/%D8%B9%D9%84%D9%88%D9%85_%D8%A7%D8%AC%D8%AA%D9%85%D8%A7%D8%B9%DB%8C) مورد استفاده قرار می‌گیرد: [گروه‌ های](http://fa.wikipedia.org/wiki/%DA%AF%D8%B1%D9%88%D9%87) محدود (مانند: قبایل و خانواده‌ها) و [طبقات اجتماعی](http://fa.wikipedia.org/w/index.php?title=%D8%B7%D8%A8%D9%82%D8%A7%D8%AA_%D8%A7%D8%AC%D8%AA%D9%85%D8%A7%D8%B9%DB%8C&action=edit&redlink=1&preload=%D8%A7%D9%84%DA%AF%D9%88:%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%D9%85%D9%82%D8%A7%D9%84%D9%87/%D8%A7%D8%B3%D8%AA%D8%AE%D9%88%D8%A7%D9%86%E2%80%8C%D8%A8%D9%86%D8%AF%DB%8C&editintro=%D8%A7%D9%84%DA%AF%D9%88:%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%D9%85%D9%82%D8%A7%D9%84%D9%87/%D8%A7%D8%AF%DB%8C%D8%AA%E2%80%8C%D9%86%D9%88%D8%AA%DB%8C%D8%B3&summary=%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF+%DB%8C%DA%A9+%D9%85%D9%82%D8%A7%D9%84%D9%87+%D9%86%D9%88+%D8%A7%D8%B2+%D8%B7%D8%B1%DB%8C%D9%82+%D8%A7%DB%8C%D8%AC%D8%A7%D8%AF%DA%AF%D8%B1&nosummary=&prefix=&minor=&create=%D8%AF%D8%B1%D8%B3%D8%AA+%DA%A9%D8%B1%D8%AF%D9%86+%D9%85%D9%82%D8%A7%D9%84%D9%87+%D8%AC%D8%AF%DB%8C%D8%AF&withJS=MediaWiki:Intro-Welcome-NewUsers.js) (مانند: جنسیّت و قومیت)

آنالیز شبکه‌های اجتماعی از طریق احکام نظری و روش‌ها و تحقیق‌های مربوط به آن از یک صنعت ضمنی به معبری تحلیلی برای پارادایم‌ها تغییر یافته‌است. برهان‌های تحلیلی از کل گرفته تا جزء؛ از ساختار گرفته تا روابط و افراد، از اخلاق گرفته تا رفتار همگی شبکه‌های سراسری را مورد بررسی قرار می‌دهند که در آن‌ها، همه رشته‌ها شامل روابط ویژه‌ای در میان جمعیت ِ تعریف شده‌اند و یا شبکه‌های فردی را مورد بررسی قرار می‌دهند که شامل رشته‌هایی است که افراد مشخصی آن‌ها را دارند از قبیل انجمن‌های خصوصی.

آنالیز شبکه‌های اجتماعی علاوه بر سروکارداشتن با اشخاص (افراد، سازمان‌ها، ایالات) به عنوان واحدهای گسسته ی تحلیل، بر روی چگونگی ساختار رشته‌ها که اشخاص و روابط میان آن‌ها را تحت تاثیر قرار می‌دهد نیز تمرکز می‌کند.



در بحث آنالیز شبکه های اجتماعی ما از بررسی پیام ها و رابطه ی بین کاربران شبکه ی اجتماعی شروع به تحقیق و بررسی می کنیم و با آنالیز داده های شبکه اجتماعی از درون مایه واقعی انبوهی از اطلاعات با خبر می شویم می توانیم پیام های شبکه های اجتماعی را آنالیز و بررسی پیشرفته بکنیم و به نحوه ی پخش شدن این پیام ها پی ببریم و با استفاده از الگوریتم های پیشرفته روی داده به آنالیز نهایی و مجموعه ای از اطلاعات و نتایج گرانبها دست پیدا کنیم.

 این کار درواقع همان کار با Big data است که در دنیا تکاپو ایجاد کرده است و حقیقتا آنالیز و بررسی اطلاعات زیاد نتایج خوب و باور نکردنی دارد. از کار هایی که پروژه های [[1]](#footnote-1)Big data کرده اند مثال های زیادی می توان زد به عنوان مثال تعداد زیادی از آمریکایی ها بر این باورند که اوباما به دلیل داشتن تیمی که کارشان انجام پروژه های Big data بود در انتخابات پیروز شد که این داستان در سایت info world نیز موجود است[[2]](#footnote-2).

این پروژه درب جدیدی را روی پروژه های Big data و هم چنین داده های شبکه های اجتماعی ایفا می کند زیرا برخلاف این پروژه ها که تنها روی محتوا بررسی انجام می دهند این پروژه روند پخش آن ها را نیز بررسی می کند.

این پروژه با پیاده سازی بر روی شبکه های اجتماعی مزیت هایی را به دنبال خواهد داشت:

1. گستره ی زیادی از کاربران را پوشش می دهد.
2. نیازی به جستجوی چندانی برای جامعه ی هدف نیست و استفاده کننده ها از پروژه مشخص است.
3. می توان گفت حدودا مرحله ی توالید و اجرایی شدن برنامه به هم متصل اند.

# **بیان مسئله**

شبکه های اجتماعی علت تعریف این پروژه هستند. فضا هایی که اطلاعات در آن به صورت شخصی یا همگانی به شیوه فوق العاده ای جا به جا می شوند. ولی چه اطلاعاتی؟ با چه مضمونی؟ به چه مخاطبی؟ در چه زمانی؟ با چه سرعتی؟ و در کل به چه شیوه ای؟

چطور یک خبر در کمتر از یک روز در تمام کشور منتشرمی شود؟ چه طور تمام دنیا یک حادثه را در کمتر از یک هفته متوجه می شود؟ اگر این اخبار غلط باشد چه؟ تشابه زیاد پیام های موجود درفضای وسیع شبکه های اجتماعی به چه علت است؟ آیا می توان از اطلاعات به نفع خود استفاده کرد؟ و این گونه مسئله آنالیز داده های شبکه اجتماعی شکل می گیرد.

البته ما شبکه های اجتماعی فراوانی داریم. پس با سوال های دیگری مواجه می شویم؛ پرمخاطب ترین شبکه اجتماعی چیست؟ محبوب‌ترین شبکه اجتماعی در ایران کدام است؟ آنالیز روی کدام شبکه اجتماعی مفید تر است؟

وایبر محبوب‌ترین و پر مخاطب‌ترین شبکه اجتماعی در ایران در سال 93 است و و جزو برترین شبکه های اجتماعی در دنیا.به گفته‌ی وزیر فرهنگ و ارشاد اسلامی[[3]](#footnote-3)، این شبکه اجتماعی بیش از 9.5 میلیون کاربر ایرانی دارد و طبق آمار الکسا[[4]](#footnote-4) بیشترین بازدید سایت وایبر توسط ایرانیان انجام شده است. (اسلامی و جنتی پنجشنبه ۱۶ بهمن 1393)

و مدتی بعد یعنی نیمه سال 94 و سال 95 محبوبترین شبکه اجتماعی در ایران تلگرام شد و تعداد کاربران زیادی از وایبر به تلگرام رفتند (این اتفاق از آنجایی شروع شد که وایبر بیش از 10 میلیون کاربر ایرانی را به خاطر سنگین شدن سرور هایش از سرور بیرون انداخت یا قسمت کمی را در اختیار آن ها قرار داد.) و تغییر اساسی ایجاد شد و هم اکنون نیز تلگرام طبق آمار الکسا در مرتبه 14 از سایت های پربازدید در ایران قرار دارد (Alexa - Top Sites in Iran 2016) و در صدر شبکه های اجتماعی و وایبر هم دیگر جزو رنک ها نمی باشد. به همین علت کار را از سر گرفتیم و شروع به پیاده سازی پروژه دیگری بر روی تلگرام کردیم که البته به نفع ما شد زیرا تلگرام اجازه دسترسی به API[[5]](#footnote-5) را می دهد.

البته محدودیت هایی هم در پروژه وجود دارد و که یکی از آن ها این است که با API تا یک سطحی می توان به اطلاعات دسترسی پیدا کرد.

اما راه حلی وجود دارد که حتی می تواند این محدودیت را رد کند و اطلاعات را در اختیار ما گزارد.

آن راه حل استفاده از روش (Robot crowling[[6]](#footnote-6)) است. این روش که از دو کلمه ی crowling به معنی خزندگی و Robot یا همان ربات تشکیل شده است. روش خزندگی که اکثرا برای جمع آوری داده ها از روی وب سایت ها استفاده می شود را با امکان ساخت ربات و استفاده از API های تلگرام ترکیب کرده و روش خزندگی بر روی شبکه های اجتماعی تلگرام با استفاده از ربات را پدید می آورد. البته قابل به ذکر است که دلیل ایججاد و کارکرد ربات های تلگرام به دور از این عمل است ولی ما با تغییر کاکرد ربات خود و توانایی های آن توانستیم بر روی ربات تلگرام همچین رباتی را پیاده سازی کنیم و حال ربات باید در شبکه اجتماعی تلگرام حرکت کند و اطلاعات و پیام ها را دریافت کند و آن را به بخش (simulator[[7]](#footnote-7)) برنامه بدهد تا آن پیام و اطلاعات وارد دیتابیس ها بشوند و مورد آنالیز و تحلیل قرار بگیرند.

لازم به ذکر است که در این پروژه خزنده و شبیه سازی به زبان C# برنامه نویسی شده است.

محدودیت دیگر این پروژه به این بر می گردد که کاربران ید موجاب به استفاده از این ربات بشوند که این کار در بخش دیگری از ربات به نام بخش طاهری انجام می شود که در شرح پروژه درباره ی آن گفته می شود.

# **پیشینه تحقیق و قوانین علمی و قوانین نظری بکار رفته**

پروژه های زیادی وجود داشته‌اند که با آنالیز انبوهی از پیام ها به اهداف خود رسیدند یا با شکست مواجه شدند که این پروژه ها به [[8]](#footnote-8)Big data معروفند و پروژه‌های Big data همچون گنجی هستند که هر چه از آن‌‌ها برداشت کنیم تمام نمی شوند. ولی این پروژه با جنبه و هدفی نو به وجود آماده است. هدف پی بردن به درون مایه پیام های پخش شده درشبکه است که در ادامه با آنالیز این پیام ها می توان به نتایجی رسید که تمام کار هایی که بقیه می خواهند انجام دهند به آسانی صورت گیرد، زیرا اگر ما بدانیم چگونه اطلاعات خود را در زمان کوتاه پخش کنیم همه کار ها آسوده می شود و در ضمن لازم به ذکر است که خلاقیت نوی این پروژه این است که تنها به پیام ها توجه نمی کند و دست به شبیه سازی و بررسی روند پخش آنها می کند. ما برای پیدا کردن توانایی در این زمینه ها مطالعات زیادی داشته ایم که برخی از مقاله ها و تحقیق های مرتبط با برنامه در انتهای مقاله و در قسمت ضمیمه ها به همراه منبع قرار گرفته اند.

# **شرح پروژه**

به دلیل محبوبیت شبکه اجتماعی تلگرام محوریت این پروژه روی این شبکه اجتماعی است.

البته این پروژه یکسال هم بر روی وایبر صورت گرفته است که محتویاتی از آن نیز در مقاله آورده شده است.

در ابتدا ربات خزنده ای (Robot crawling) طراحی شده است که با جریان در شبکه اجتماعی تلگرام می تواند دائما پیام ها ، گروه ها و افراد مختلفی را بیابد و مورد آنالیز قرار دهد.

البته وایبر هم شیوه ی بدست آوردن داده های خود را دارد ولی ضعیف تر از تلگرام است.

و قسمت اصلی دیگر شبیه ساز (simulator) نام دارد که هم در قسمت و وایبر و هم در قسمت تلگرام وجود دارد.

که البته در قسمت تلگرام خیلی قوی تر از قسمت وایبر کار می کند.

## **شرح قسمت وایبر پروژه(قسمت فرعی)**

### **بدست آوردن داده ها**

هر اکانت می تواند نسخه ی پشتیبانی(Backup) اطلاعات خود را از برنامه خارج کرده وآن ها را داشته باشد حال اگر بکاپ تنها 100 نفر را داشته باشیم می توانیم پیام های مکالمات هزاران نفر را رصد کنیم و روی آنها آنالیز و بررسی انجام دهیم. که این کار به وسیله ی این شبیه ساز انجام می شود.

اما جمع آوردی این داده ها به چه شکلی است؟

جمع آوری داده ها به چندین شکل صورت می گیرد از جمله:

1. به صورت مصاحبه با افراد و گرفتن فایل پشتیبانی وایبرشان
2. استفاده از برنامه محافطت از حریم خصوصی: جهت دریافت فایل از افرادی که به حریم خصوصی خود حساس اند.
3. استفاده از سایت: توانایی ارسال نسخه ی پشتیبانی به سایت و توانایی ترقیب افراد توسط مطالب
4. استفاده از برنامه های جاسوسی(Spyware): اتصال به برنامه ی جاسوسی پیام ها و گرفتن پیام های وایبر اشخاص

بعد از گرفتن داده ها برای به ازای هر گروه وایبر فایلی بی نظم و ترتیب و بدون قالب بندی سالم و با فرمت CSV که شامل اطلاعات ذیل می باشدگرفته می شود:

* گروه های هر اکانت
* مخاطبان فرد(اسم و شماره تلفن)
* متن پیام ها
* زمان پیام ها(ساعت و تاریخ پیام ها)
* فرستنده پیام ها

**حریم خصوصی**

اولین کاری که این برنامه انجام می دهد خواندن این فایل ها می باشد اما یک مشکل در این بین وجو دارد که آن هم پیام های خصوصی افراد است.

پیام های خصوصی افراد باید قبل از ورود به شبیه ساز حذف شوند که هم روند آنالیز داده های مهم را کند نکنند و هم امنیت اطلاعات افراد بر قرار باشد، این کار به روش زیر صورت می گیرد:

* از طریق دستوراتی که به شبیه ساز گفته می شود تا قبل از ورود نسخه پشتیبانی به شبیه ساز آن دستورات را روی فایل ها انجام دهد.

بعد از اینکه برنامه فایل هارا خواند و کارهایی که گفته شد صورت گرفت تمام گروه ها در لیستی قرار می گیرند و تمام اطلاعات فایل ها به صورت جدول بندی شده و منظم در یک ِDataGridView نمایش داده می شود.

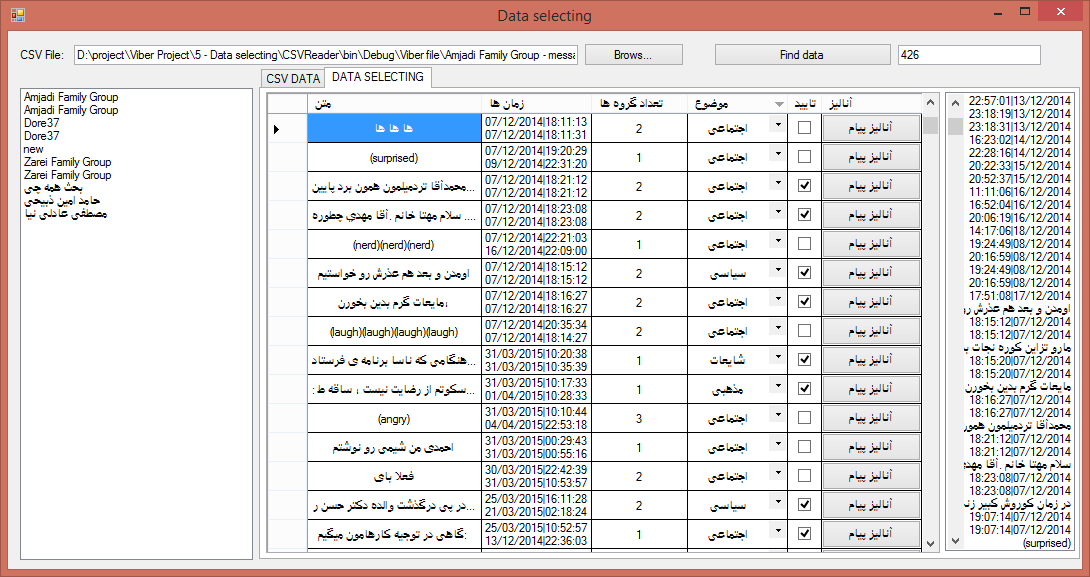
که در هر رکورد آن می توان نام فرستنده، زمان پیام، گروه انتشار پیام و در آخر متن پیام را مشاهده کرد.

### **بخش پالایش محتوا پیام**

بعد از خواندن و جدول بندی کردن اطلاعات نوبت حذف فایل های بیهوده یا همان junk است.

که این کار با استفاده از جستجو در پیام ها و استفاده از کلمات کلیدی و کلمات به کار برده توسط خود وایبر است به طور مثال نام شکلک هایی که در وایبر استفاده می شود.(مثال: (smiley)(sad)(wink)) این نام ها در این قسمت شناسایی و بررسی می شوند و اکثرا به عنوان عبارات بیهوده حذف می شوند. این قسمت خود یکی از مهمترین بخش‌ها است چراکه باعث بهینه سازی و تسریع عملکرد برنامه می شود.

### **یافتن تشابه و نمایش صفات**

در این قسمت با توجه به میزان تشابه پیام ها به یک دیگر و با استفاده از الگوریتم های پیشرفته و حتی هوش مصنوعی داده‌ها انتخاب شده و سپس پر تکرار ترین پیام ها گلچین می شوند و در این مرحله نیز باید پیام ها یکبار دیگر مرور شوند و پیام های بیهوده حذف شوند.

حال ما محبوبترین و پر تکرارترین و بهترین و مخاطبپسند ترین پیام ها را در دسترس داریم با تمام مشخصات، توضیحات و جزئیات آن

در این قسمت برنامه با استفاده از الگوریتم های مانند [[9]](#footnote-9)weight word موضوع و محتوای پیام هارا بررسی می کند و آنها را دسته بندی می کند و با استفاده از تاریخ و ساعت ها و دقیقه ها زمان پخش آن هارا مشخص می کند و سرعت و نحوه پخش آن ها را آنالیز می کند و در نمودار نمایش می دهد. نمونه نمودار :

و بعد از آنالیز آن ها با استفاده از الگو برداری از آن پیام ها می توان اطلاعات مفید و باارزش بسیاری بدست آورد برای مثال می توان فهمید چگونه می توان یک پیام را برای یکسری مخاطبان خاص دریک زمان معین ارسال کرد و آنها را به توجه واداشت و یا اینکه درصد نسبی سرعت انتشار داده هارا در موضوعات مختلف برای مخاطبین مختلف فهمید و با استفاده از این پیام های تبیلغاتی را پیکربندی و ارسال کرد و نتیجه گرفت.

## **شرح پروژه قسمت تلگرام(قسمت اصلی)**

### **ربات خزنده(Robot crawling)**

در پروژه برای بدست آوردن داده ها از یک ربات خزنده استفاده می شود این ربات از دو بخش تشکیل می شود بخش ظاهری و بخش باطنی

### **بخش ظاهری:**

ربات تلگرام طراحی شده در بخش بیرونی یک شایعه یاب است که جدا از کار آنالیز عمل می کند شیوه ی عملکرد او به این شکل است که به بزرگترین دیتابیس های شایعات و جواب( (شایعات) (ویکی هوآکس)) [[10]](#footnote-10)آن ها متصل می شود و در صورت گفته شدن شایعه به اعضای گروه اطلاع می دهد و جواب شایعه را نیز بیان می کند. با این ویژگی جالب و منحصر به فرد که خود می تواند یک پروژه جالب باشد کاربران را به استفاده و پخش این ربات مشتاق می کند. و کاربران درصدد استفاده از این ربات قرار می گیرند تا از وجود شایعات و پخش آنان جلوگیری کنند و همچنین قدمی در جلوگیری از پخش شایعات بر داشت.

### **بخش باطنی:**

در این بخش **ربات خزنده** جدا از کار ظاهری شروع به خزیدن در شبکه های اجتماعی می کند و ریز به ریز اتفاقات را در برنامه مطابق نمونه پایین اعلام می کند. این کار در حالت عادی مشکل است چراکه بات های تلگرام برای این عملکرد تعریف نشده اند و محدودیت دارند برای همین ما در برنامه نویسی خود درخواست های خود را به تلگرام برای دریافت اطلاعات در قالب گره هایی(Threads[[11]](#footnote-11)) که به صورت موازی اطلاعات را از سرور تلگرام در خواست می کنند(منظور همان Request های ربات به تلگرام است )اجرا می کند با این روش می توان محدودیت ربات را تا حد ممکن بالا برد. سپس این پیام هارا دریافت کرده و به دیتابیس[[12]](#footnote-12) ارسال می کند سپس اطلاعات ارسال شده به دیتابیس جدول بندی می شود. این جداول به صورت دسته بندی شده و کامل اطلاعات را در بر می گیرد. نکته مثبت اینکار این است که این ربات و تمام دیتابیس ها در سرور ما قرار دارند و این کار دسترسی به تمام اطلاعات را برای ما سهل و آسان می گرداند.

**حریم خصوصی**

در این قسمت پیام های خصوصی دخیل نمی شوند یا به میزان خیلی کم دخیل می شوند چراکه ربات فقط پیام های گروه ها را مورد بررسی قرار می دهد نه پیام های بین دو نفر یا پی وی هارا

### **پالایش محتوای پیام**

حال در پیام های گرفته شده شروع به حذف داده های بیهوده می کنیم از جمله شکلک ها و حرف هایی که به درد تحلیل نمی خورد. این حرف هارا با استفاده از دیتابیسی که تهیه شده حذف می کنیم که نام آن دیتابیس در پروژه به هجویات است و از تداخل دیتابیس های موضوعی بدست می آید که کاملا خودکار کار میکند و در ادامه درباره ی آن گفته خواهد شد. این دیتابیس از تعداد تکرار کلمات از دیتابیس موضوعی یک بایگانی درست می کند و با استفاده از آن می توان کلمات هجو را شناسایی کند و درصورت مشاهده آن هارا حذف کند.

### **یافتن تشابه و نمایش صفات**

در این قسمت دیتابیس گرفته شده از ربات خزنده در حالی که هجویاتش حذف شده بازبینی می شود و پیام هایی که بیش تر پخش شده اند با مشخصاتشان جدا شده و در دیتابیس دیگری جای می گیرند البته در این پخش شدن فقط به تکرار عینی پیام بر نمی گردد بلکه با استفاده از الگوریتم های تشابه یابی[[13]](#footnote-13) این کار صورت می گیرد و البته قسمت تشابه یابی این پروژه در تلگرام بسیار قوی شده است و از الگوریتم خوبی بهره می برد که تمام نکات را برای مقایسه دو رشته مورد توجه قرار می دهد از جمله کلمات مشابه و فاصله ی کلمات از کلمات مشابه مکان کلمات مشابه تعداد کاراکتر ها و البته سرعت بسیار بالایی هم دارد که بیش از 50 پیام در ثانیه است.

این دیتابیس جدید شامل متن پیام ها، تعداد تکرار پیام، تاریخ پخش آن ها و آیدی ارسال کننده گان و گروه ارسال کننده و موضوع پیام می باشد.

ما می دانیم که با استفاده از ربات خزنده می توانیم به متن پیام و تعداد تکرار آن و تاریخ پخش و ما بقی اطلاعات به جز موضوع پیام دسترسی پیدا کنیم ولی موضوع را از کجا می توان یافت.

### **ایجاد دیتابیس های موضوعات مختلف**

این پروژه که درسال 94 در قسمت وایبر مانور کمی در این بخش داده بود امسال با مانور بسیار وارد این قسمت شد چراکه یکی از نقص های بزرگ این پروژه که در جشنواره های سال قبل به او وارد شد این بخش بود.

در سال گذشته بنای این پروژه اینگونه بود که باید دیتابیس کلمات موضوعی را از برچسب های سایت ها با موضوعات مختلف به صورت دستی و جداگانه وارد دیتابیس کرد خوب اینکار معایبی داشت ازجمله:

* از جمله دستی بودن و گرفتن وقت بسیار زیاد که شاید گذاشتن این وقت ممکن نبود.
* بدست آوردن تعداد کم و محدودی از کلامات که نمی توانست جامعیت پیدا کند.
* به روز نبودن و عدم وجود بعضی از کلمات روز

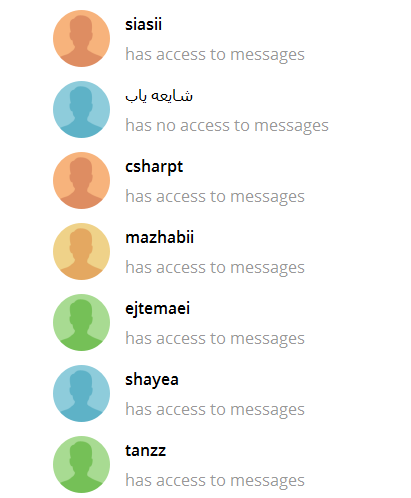
اما امسال در بخش شبکه اجتماعی تلگرام انواع ایده را برای این موضوع چک کردیم تا درنهایت به این نتیجه رسیده شد که دیتابیس موضوعات باید چند ویژگی داشته باشد:

* دستی نباشد و به صورت اتوماتیک باشد
* همیشه به روز باشد و از جدیدترین لغات روز بر خوردار باشد
* محدودیت زیادی نداشته باشد و طوری تامین شود که بتواند جامعیت داشته باشد

از موارد بالا دریافته شد که پیش نیاز یک همچین دیتابیسی برای قدم اول آنلالین بودن آن است که چون ما در پروژه ی تلگرام بر خلاف وایبر که آفلاین کار می کرد آنلاین کار می کنیم خیالمان از این بابت راحت است.

بعد به این نتیجه رسیده شد که تنها استفاده از برچسب سایت ها جامعیت و بروز بودن ایجاد نمی کند چرا که خیلی از سایت های بزرگ برچسب گزاری نمی کنند یا برچسب های خود در معرض دید نمی گزارند و بدست آوردن آنلاین برچسب ها و آن هم برچسب های بروز امکان پذیر نیست.

پس به سراغ منبع دیگری باید رفت، بعد از آن موارد مختلفی بررسی شد از جمله کانال های موضوعی تلگرام یا اکانت های اینستاگرام ولی در همه ی موارد یک محدودیت در محتوا یا پیاده سازی وجود داشت تا اینکه به استفاده از RSS منتهی شد ولی آر اس اس چه سایتی بدیهی است که سایت های خبری بروزترین سایت های کشور می باشند و در صورت اتصال به بهترین سایت های خبری می توان در زمان نه چندان زیادی دیتابیس جامعی به دست آورد. پس برای این قسمت به این نتیجه رسیده شد که از آر اس اس بزرگترین سایت های خبر گزاری کشور( (خبر گزاری فارس) (خبر گزاری تابناک) (خبر گزاری ایسنا) (خبر قطعه) (خبر گزاری پارسیک) (خبر گزاری خبر فارسی) (خبر گزاری شهر خبر)) استفاده کرد و اینکار نکته مثبتی هم داشت و آن این بود که این سایت ها خود اخبار را بر اساس موضوعات دسته بندی می کنند و این قابلیت ایجاد می شد که دیتابیس هر موضوع را از همان اخبار گرفت.



حال که منبع معلوم شد باید برای بدست آوردن آن ها برنامه ریخته می شد ابتدا به عنوان اولین ایده از ربات های تلگرام برای بدست آوردن دیتا های دیتابیس موضوعی استفاده شد.

ربات خزنده به آن ها متصل می شد و دیتابیس موضوعی را از آن ها تهیه می کرد ولی در این ایده با مشکلی مواجه شد و معلق ماند ولی در ایده بعد در خود برنامه از کلاس ها برای بدست آوردن دیتابیس استفاده شد و به نتیجه رسید.سپس مانند قسمت وایبر از الگوریتم weight word برای وزن دهی به کلمات استفاده می شد ولی باید در نظر داشت که این دیتابیس ها با این روش لغات مشابهی پیدا می کنند که در حقیقت جزو هیچ کدام دسته نیست مانند من، دیروز، بگو یا ... به همین علت کدی نوشته شد که اگر کلمه ای به تعداد زیاد در دیتابیس های مختلف تکرار شد جزو هجویات به شمار برود. ولی باز هم نمی توان گفت که متنی صد در صد یک موضوع می گیرد ولی درصد موضوعی با اختلاف زیاد بیشتر از بقیه موضوعات می شود که آن موضوع متن شناخته می شود. به عنوان مثال متن مربوط به نمودار زیر سیاسی است:

در داوری سال گذشته خوارزمی این ایراد به پروژه وارد شد که الگوریتم weight word قابلیت موضوع بندی ضرب المثل را ندارد طبق کار هایی که روی این ایراد انجام گردید به این نتیجه رسیده شد که اولا در روند آنالیز این اتفاق هیچ خللی وارد نمی کند ولی برای مطمئن شدن ما می توانیم از یک تکنیک هوش مصنوعی استفاده کنیم این تکنیک در جاهای مختلفی از جمله موتور جستجو گر گوگل کاربرد دارد به این شکل عمل می کند که با استفاده از داده های قبلی می تواند داده ی کنونی را به یک جهت متمایل کند البته باید خود داده نیز توانایی تمایل را داشته باشد و از داده های دیگر پرت نباشد که این تکنیک در این برنامه به این گونه عمل می کند که پیام های قبلی و بعدی آن پیامی که موضوع بندی نشده توجه می شود و اگر خود آن پیام پتانسیل تمایل به موضوعات آن ها را داشت به آن موضوع مرتبط می شود. این پتانسیل پیام به این بر می گردد که اون پیام در جریان یک بحث گفته شده باشد یا موضوع کنونی به چیز دیگری بر نگردد.

### **آنالیز و تحلیل ها**

**حال که ما جدولی با تمام مشخصات پیام و حتی موضوع او داریم می توانیم دست به هر تحلیلی که خواستیم بزنیم.**

**یکی از مهم ترین تحلیل ها مشخص کردن سرعت پیام ها است برای بدست آوردن سرعت باید زمان پخش یک پیام و گروه هایی منتشر شده است را با استفاده از جدول پیدا کنیم مثلا در این ماه چگونه و در چند گروه منتشر شده است و همچنین از روند پخش آن استفاده کنیم. یک نمونه نمودار روند پخش پیام:**

(دو نمونه آنالیز گفته شده در قسمت بالا جزو آنالیر های هر پیام می باشد که در قسمت نمودار های پروژه در باره ی آنها به همراه تصویرشان از پروژه اطلاعات کاملی گفته می شود.)

آنالیز های مهم دیگری که می توان گفت پخش یک پیام است یعنی ما با اطلاعات بدست آمده از گروه ها و با بررسی سر منشا موضوعات هر پیام بتوانیم تشخیص بدهیم که چگونه می توان یک پیام را با ساختار های مشخص و جامعه ی هدف معلوم پخش کرد و با روند تقریبا کنترل شده ای حرکت پیام هارا حدس زد.

که برای اینکار ساختار پیام هارا مد نظر قرار می دهیم و پیام های گروه هارا بررسی می کنیم مثلا چه لغاتی در یک پیام بررسی شده و در چه فرمتی چه گروهی از این نوع پیام فراوان است و چه افرادی در آن گروه هستند و آن گروه با چه گروه ها و افرادی در ارتباط است.

آنالیز دیگری اینگونه است که می توانیم دربازه های زمانی مشخص شده ای با استفاده از موضوع و تعداد تکرار پیام میزان پیام های پخش شده ی هر موضوع را در جامعه ی هدف مورد نطر دریابیم و از آن به بهترین نحو استفاده کنیم. به عنوان مثال اگر در ماهی پیام های سیاسی بیشتر باشد یعنی رویداد مهم یا تحول سیاسی ای رخ داده است که در جامعه ی هدف تاثیر زیادی گذاشته است یا مثلا می توان فهمید که یک شایعه تا چه مقدار فراگیر شده است و ... یک نمونه نمودار که دریکی از ماه هاپیام های شایعه از ماه های دیگر بیشتر پخش شده است:

### **شبیه سازی**

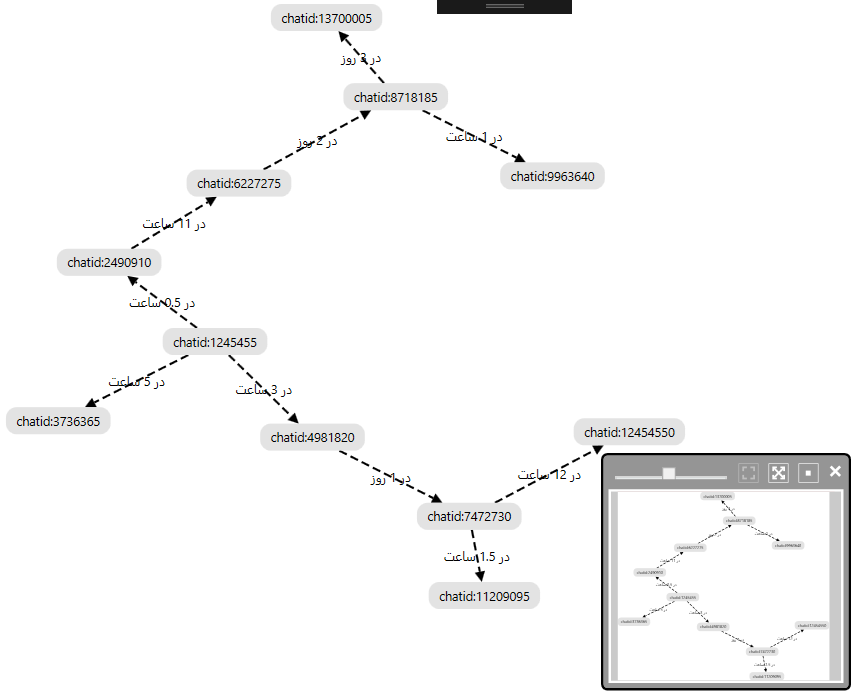
همان طور که از نام پروژه بر می آید(آنالیز انتشار داده ها در شبکه های اجتماعی) ای پروژه با انتشار داده ها سر و کار دارد و روند پخش آن ها را مورد بررسی قرار می دهد پس در قسمت شبیه سازی هم روند پخش پیام شبیه سازی می گردد.

در این بخش با استفاده از زمان متوسط پخش پیام و زمان پیام ها در گروه های مختلف و همچنین رابطه ی گروه ها با یکدیگر و افراد ارسال کننده ی مستقیم و غیر مستقیم پیام روند انتشار پیام تا حد نسبتا دقیق دریافت می گردد و شبیه سازی می شود.

در این بخش علاوه بر نمایش مکان هایی که پیام مشاهده شده است مکان های دیگری که توسط شبیه ساز نیز حدس زده شده است نمایش داده می شود که این حدس ها با بهره گیری از الگوریتم هایی که برای شبیه ساز نوشته شده است صورت می گیرد و این الگوریتم ها مشخصات مذکور و روابط بین آن ها استفاده می کند.

و در نهایت نیز این شبیه سازی در قالب یک گراف جهت دار نمایش داده می شود.

**نمونه گرافی از پخش یک پیام:**



### **کنترل جامعه هدف**

در این بخش می توان پیام های یک جامعه ی هدف خاص را کنترل کرد. این کنترل شما چند ویژگی می باشد:

1. ایجاد کلمات خاص که در صورت مشاهده ی کلمات هشدار می دهد.
2. انتخاب موضوع خاص که درصورت گفتن پیامی با آن موضوع برنامه اطلاع می دهد.
3. یا می توان این محدودیت ها را تنها برای کاربران خاصی قرار داد.

و در آخر نیز می توان پیامی مشخص کرد که در صورت به وجود آمدن شرایط تعیین شده آن پیام به جامعه ی هدف ارسال شود.

### **کنترل تقریبی و ارسال پیام**

در این پروژه بعد قوی تر شده دیتابیس ها این امکان وجود دارد که بتوان به هر کجا که ربات در آن قرار دارد از طرف ربات پیام داده بشود.

ولی ما برای پخش پیام نیازی نیست به همه گروه ها پیام دهیم چرا که این کار ارزش پیام را کاسته و توجهات را از آن کم می کند.

به همین دلیل در این برنامه با گرفتن کلمات کلیدی پیام مورد نظر و موضوع آن لیست بهترین گروه ها برای پخش پیام گفته می شود و همچنین در حین نوشتن پیام کلماتی که طبق پیام های گذشته برای گروه های مورد نظر معروف بوده اند پیشنهاد داده می شود.

### **بات سواری**

این قسمت تعبیه شده تا بتوان از داخل خود برنامه با یک گروه از طریق ربات حرف زد و پیام های آن گروه را مشاهده کرد.

### **سرور اجرایی ربات**

برای اجرای این ربات ما نیاز به سرور داریم برای این سرور یک محدودیت هایی را باید در نظر می گرفتیم:

باید قابل اطمینان بوده و داخلی باشد

باید به ما رم بالا و پهنای باند بالایی بدهد

کم هزینه باشد

پتانسیل اجرای برنامه های درخواستی ما را داشته باشد

اما بدست آوردن همچین سروری راحت نیست چراکه طراحان این برنامه به دلیل سن پایین توسط هر میزبانی تایید نمی شوند و همچنین پیاده سازی این مشخصات در کنار هم کار ساده ای نیست و همچنین به دلیل طرح کلی این برنامه امنیت اطلاعات خیلی اهمیت دارد به همین دلیل ما برای اجرای برنامه خودمان شروع به ایجاد سرور کردیم و با نصب sqlserver بر روی آن و سایر برنامه های لازم و همچنین اتصال آن به یک شبکه اینترنت با پهنای باند بالا و همچنین ایجاد این توانایی که بتوان از هر مکان به آن وصل شد خودمان سرور خودمان را تامین کردیم و هزینه ی کمتری هم برایمان داشت و در عین حال کارکرد بیشتری

# **چشم انداز توسعه**

**شاید برای همچین برنامه هایی نتوان هیچ انتهایی را به طور کامل تعیین کرد ولی ما برآنیم تا جای ممکن این پروژه را جلو ببریم.**

**یافتن سر منشا های موضوعی:**

**یکی از کار هایی که باید در آینده نزدیک در این پروژه انجام شود شناسایی گروه هایی که اکثرا درباره ی یک موضوع خاص حرف می زنند است و سپس فهمیدن ارتباطی که بین آنها وجود دارد چراکه با این کار می توانیم از منشا های بزرگ یک موضوع را پیدا کنیم و بتوانیم پیام های خود را که با آن ساختار است از آنجا پخش کنیم.**

**عضویت ربات در گروه های بیشتر:**

**گرچه ربات خزنده هم اکنون در تعداد زیادی گروه پرجمعیت عضو است که بعضا تعداد افراد آن ها به بیش از هزار نفر می رسد و همچنین آن ها هم برای استفاده از ویژگی شایعه یابی ربات آن را پخش می کنند ولی می توان ایده های بیشتری را برای پخش آن پیاده سازی کرد یکی از این ایده ها نوشتن برنامه اندرویدی جهت معرفی کردن ربات است و همچنین یکی دیگر از ایده های جالب و نه چندان سخت ایجاد کانال یا سایتی است که اطلاعات جزیی از آنالیز را در معرض دید می گزارد ولی تنها کسانی که با این ربات در گروهایشان در ارتباطند قابلیت عضو شدن در این کانال را داشته باشند و اینکار کار راحتی است و شاید بزودی انجام گیرد چون هیچ بخش دستی ای نخواهد داشت و حتی مدیریت آن کانال هم بر عهده ی ربات خزنده می شود.**

**اضافه شدن قسمت شخصیت شناسی به پروژه:**

**برای اینکار ربات دیگری به پروژه اضافه می شود که در قسمت ظاهری خود اطلاعات مورد علاقه کاربران را بر اساس تاریخ دست بندی می کند و در صورت نیاز کاربر آن هارا بازیابی نیز خواهد کرد و اینگونه می تواند با آنالیز آن پیام ها شخصیت آن کاربر را بشناسد و از وجود او در گروه ها درباره گروه اطلاعات کسب کند و یا اینکه به کاربر گروه هایی مورد علاقه و مطابق شخصیتش معرفی کند.**

# **کاربرد طرح**

1. **یافتن سر منشا پخش پیام ها با ساختار و جامعه ی مورد نظر و حدس تقریبا دقیق روند پخش پیام**
2. **کنترل و بررسی پیام های یک جامعه ی هدف**
3. **کنترل چارچوب یک جامعه و مجموعه پیام**
4. **هدف های تجاری و فروش آنالیز های بدست آمده و پخش پیام**
5. **یافتن رابطه افراد مختلف با یک دیگر اگر چه این موضوع جزو اهداف فرعی به شما می آیند**
6. **استفاده از بخش ظاهری ربات و جلوگیری از پخش شایعات در تلگرام**

**. کاربرد های بسیاری که می تواند یک آنالیز گز حرفه ای و هوشمند داشته باشد.**

**.**

**.**

# **مزایای پروژه**

* آنالیز متفاوت شبکه های اجتماعی نسبت به آنالیز های دیگر
* کار بر روی پر طرفدارترین شبکه ی اجتماعی کنونی در ایران (تلگرام)
* کامپیوتری بودن پروژه و دستی نبودن هیچ کدام از مراحل
* کار با حجم عظیمی از اطلاعات(Big data)
* این پروژه شاید در نگاه اول دیتابیس ضعیفی داشته باشد ولی چیزی که این دیتابیس را خاص می کند زمان است دیتابیس این پروژه چه دیتابیس های موضوعی چه دیتابیس های خود پیام نه با گذر زمان کهنه نمی شوند بلکه در طول زمان با کار بر روی داده های جدید سیر تکاملی پیدا کرده و رشد می کنند و در نهایت کامل ترین و جامعترین دیتابیس در این موضوعات خواهند شد.
* نوع برنامه نویسی سعی شده تا حد ممکن بهینه باشد و ربات خزنده هم توانایی دریافت بیش از 50 پیام در ثانیه را داراست.

# **معایب پروژه**

* به دلیل آنلاین بودن پروژه تلگرام، این برنامه نیاز به اجرا بر روی سرور دارد و با توجه به تکامل خود به خودی پروژه در طول زمان نیاز به یک سرور نسبتا قوی پیدا می کند.
* ما معتقدیم که با حجم عظیمی از مطالب بدست آمده کار های زیادی می توان انجام داد ولی کار هایی که ما بر روی این داده ها انجام می دهیم کافی نیست زیرا از دریای گنج برداری کم است.

**منابع**

بدون تاريخ. http://papro-co.ir/entity-framework-%DA%86%DB%8C%D8%B3%D8%AA%D8%9F/.

بدون تاريخ. https://fa.wikipedia.org/wiki/%DA%AF%D8%B1%D8%A7%D9%81\_%D8%AC%D9%87%D8%AA%E2%80%8C%D8%AF%D8%A7%D8%B1.

2016. *Alexa - Top Sites in Iran.* 25 4. http://www.alexa.com/topsites/countries/IR.

develope, انجمن سایت مرکز. بدون تاريخ. *Big Data.* http://www.developercenter.ir/.

Halleux, Jonathan de. 2007. 23 Apr . http://www.codeproject.com/Articles/5603/QuickGraph-A-100-C-graph-library-with-Graphviz-Sup#xx5110839xx.

2014. *http://stackoverflow.com.*

2015. *www.codeproject.com.*

اسلامی, وزیر فرهنگ و ارشاد, و علی جنتی. پنجشنبه ۱۶ بهمن 1393. *تعداد کاربران وایبر.* فارسی, دویچه‌وله.

پدیا, ویکی. بدون تاريخ. *String-to-string\_correction\_problem.* https://en.wikipedia.org/wiki/String-to-string\_correction\_problem.

تبیان, سایت. بدون تاريخ. *مفهوم و معنی RSS.* http://www.tebyan.net/newindex.aspx?pid=29963.

تلگرام, سایت رسمی. بدون تاريخ. *ربات های تلگرام.* https://core.telegram.org/.

بدون تاريخ. *خبر قطعه.* http://isna.ir/page/news-feed.xhtml?lang=fa.

بدون تاريخ. *خبر گزاری ایسنا.* http://isna.ir/page/news-feed.xhtml?lang=fa.

بدون تاريخ. *خبر گزاری پارسیک.* http://www.parseek.com/About/NewsRssFeeds.aspx.

بدون تاريخ. *خبر گزاری تابناک.* http://www.tabnak.ir/fa/rss.

بدون تاريخ. *خبر گزاری خبر فارسی.* http://khabarfarsi.com/%D8%A7%D8%AE%D8%A8%D8%A7%D8%B1\_%D8%B3%DB%8C%D8%A7%D8%B3%DB%8C.

بدون تاريخ. *خبر گزاری شهر خبر.* http://www.shahrekhabar.com/rssfeeds.

بدون تاريخ. *خبر گزاری فارس.* http://www.farsnews.com/RSSLinks.

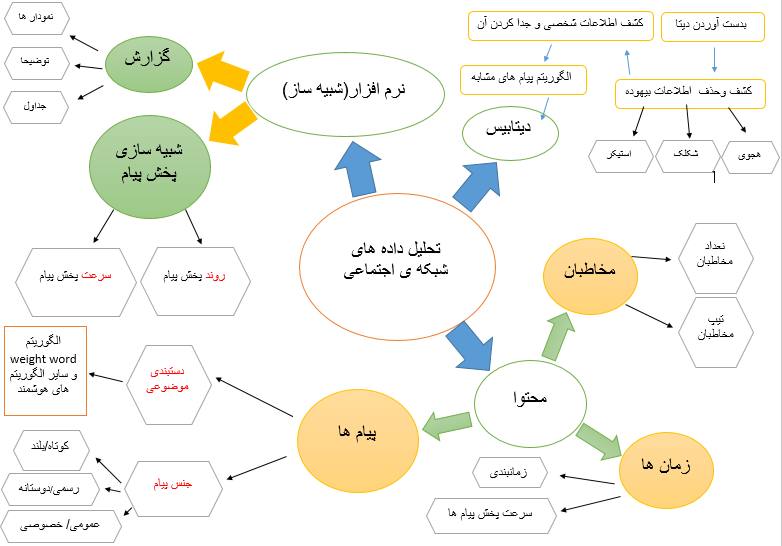
بدون تاريخ. *ربات تلگرام.* http://websima.com/%D8%B1%D8%A8%D8%A7%D8%AA-%D8%AA%D9%84%DA%AF%D8%B1%D8%A7%D9%85.

بدون تاريخ. *شایعات.* http://shayeaat.ir/.

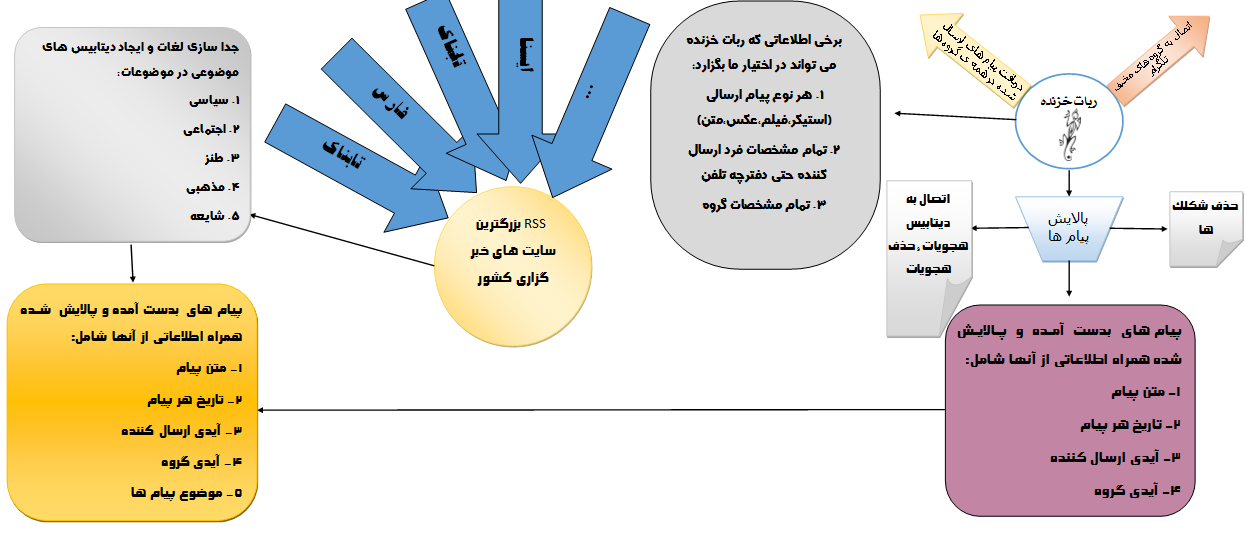
بدون تاريخ. *ویکی هوآکس.* http://wikihoax.org/.

# **نمودار ها**

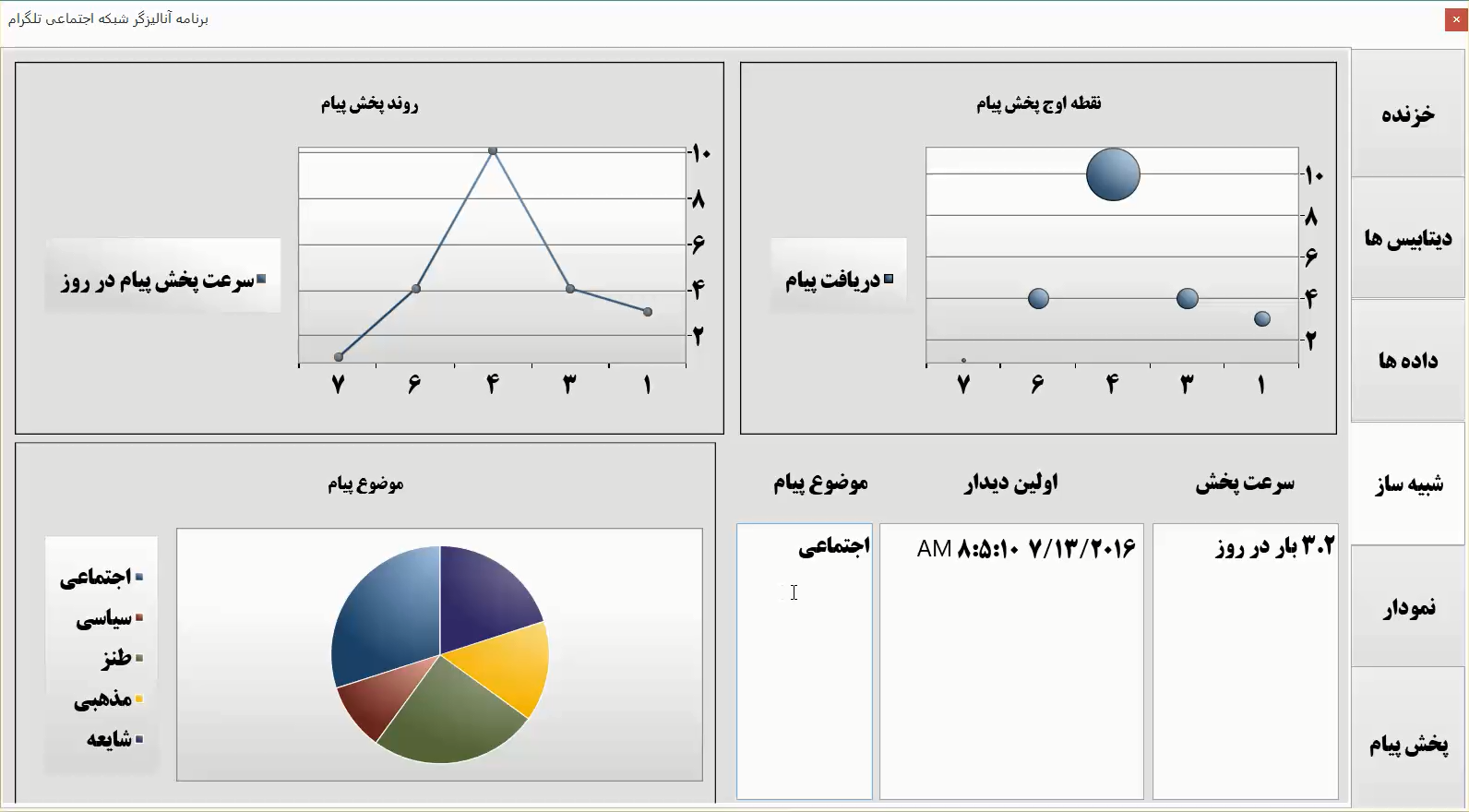
## نمودار مولفه های پروژه



## نمودار بدست آوردن دیتا در تلگرام



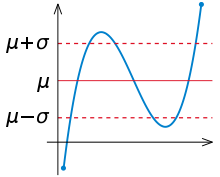
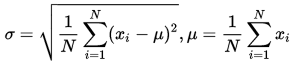
## نمونه ای از بخش نمودار های هر پیام در پروژه



**نمودار موضوع پیام:**

در برنامه با استفاده از دیتابیس موضوعی کلمات که در بطن مقاله از آن نام برده شده است می توان پیام هارا موضوع بندی کرد و این نمودار دایره ای نیز بر اساس ارتباط پیام با آن موضوعات رسم شده است.

**نمودار سرعت و روند پخش پیام:**

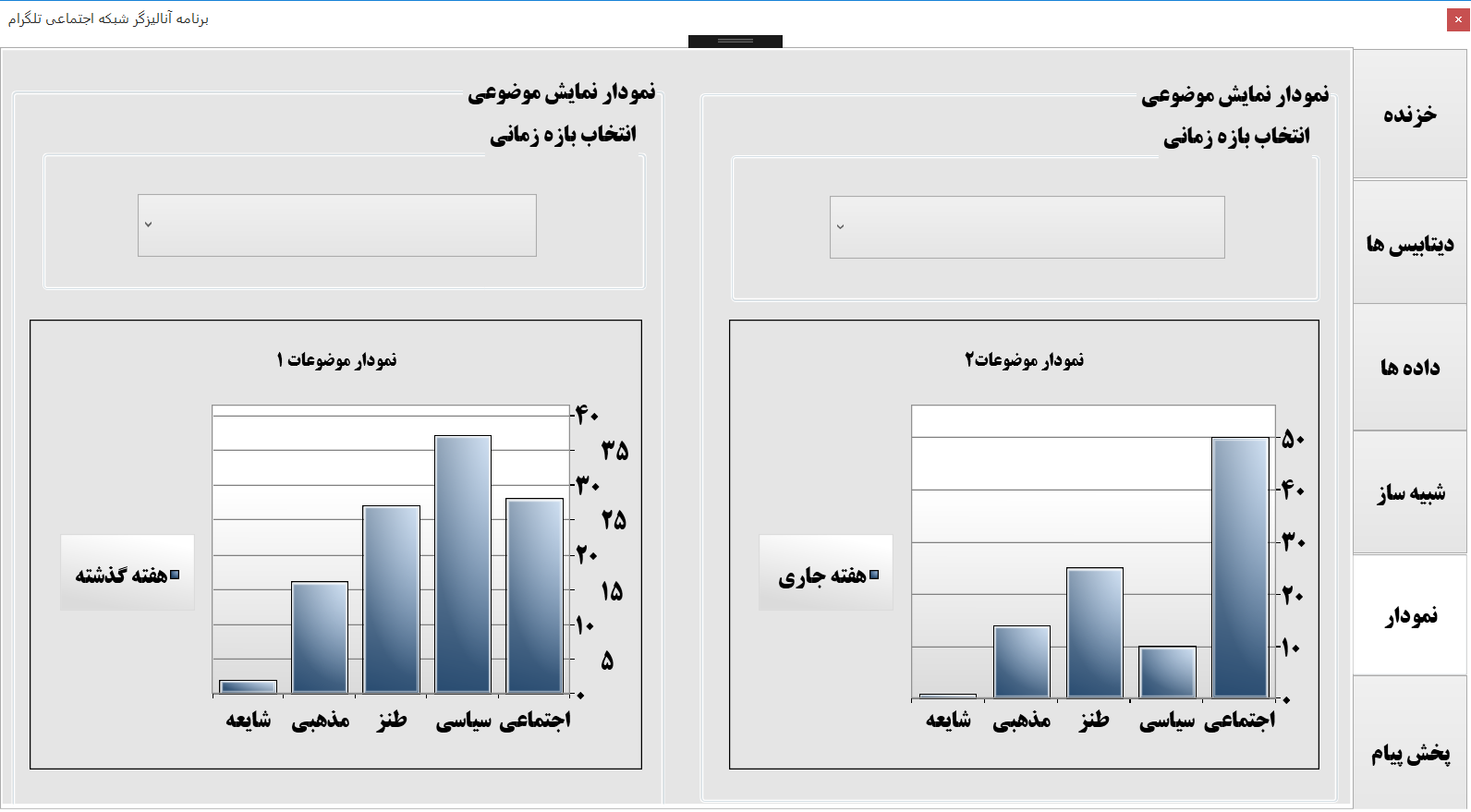
با استفاده از زمان های پخش پیام های مشابه[[14]](#footnote-14) و تعداد آن ها و استفاده از فرمول و مفهوم واریانس(Variance) و انحراف معیار(Standard Deviation) این نمودار رسم می شود و از روی همین نمودار هم می توان میانگین سرعت پخش پیام هارا نیز متوجه شد.

**نمودار نقطه اوج پیام:**

از روی این نمودار می توان بازه زمانی که پیام در آن زمان بیشترین پخش را داشته است مشاهده کرد.

به عنوان مثال اگر پیام ها در مورد ولادت ها یا یک رویداد سیاسی یا اجتماعی باشد در نقطه اوج پیام ها همان روز ولادت یا روز دقیق آن رویداد را نشان می دهد.

## نمونه نمودار های کلی پیام ها در پروژه

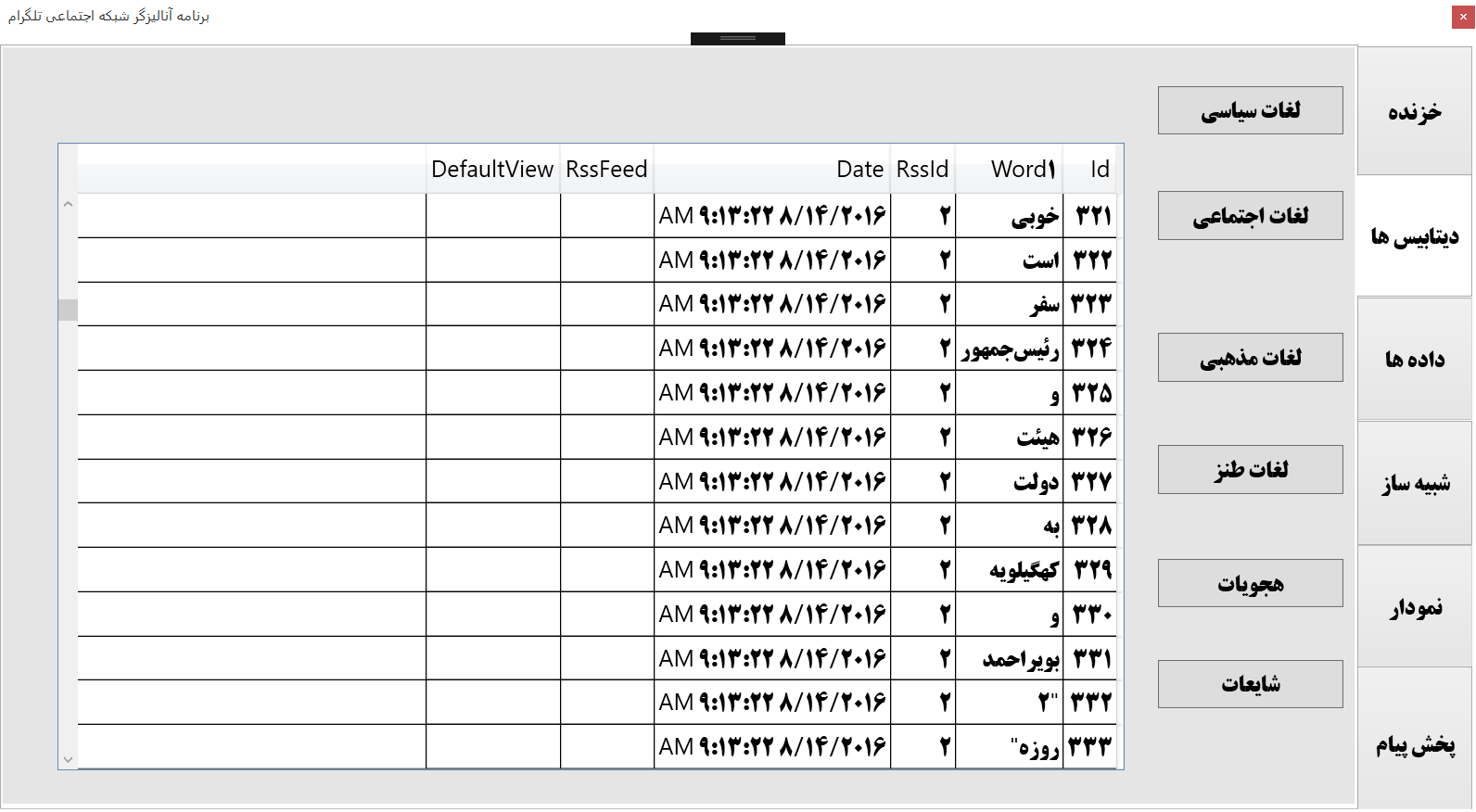


این تصویر از برنامه به مقایسه کلی پیام ها در بازه های زمانی مشخص بر می گردد و به شناسایی اتفاق و رویدادی در بازه زمانی مشخص می انجامد.

# **ضمایم پروژه:**

## عکس هایی از پروژه:

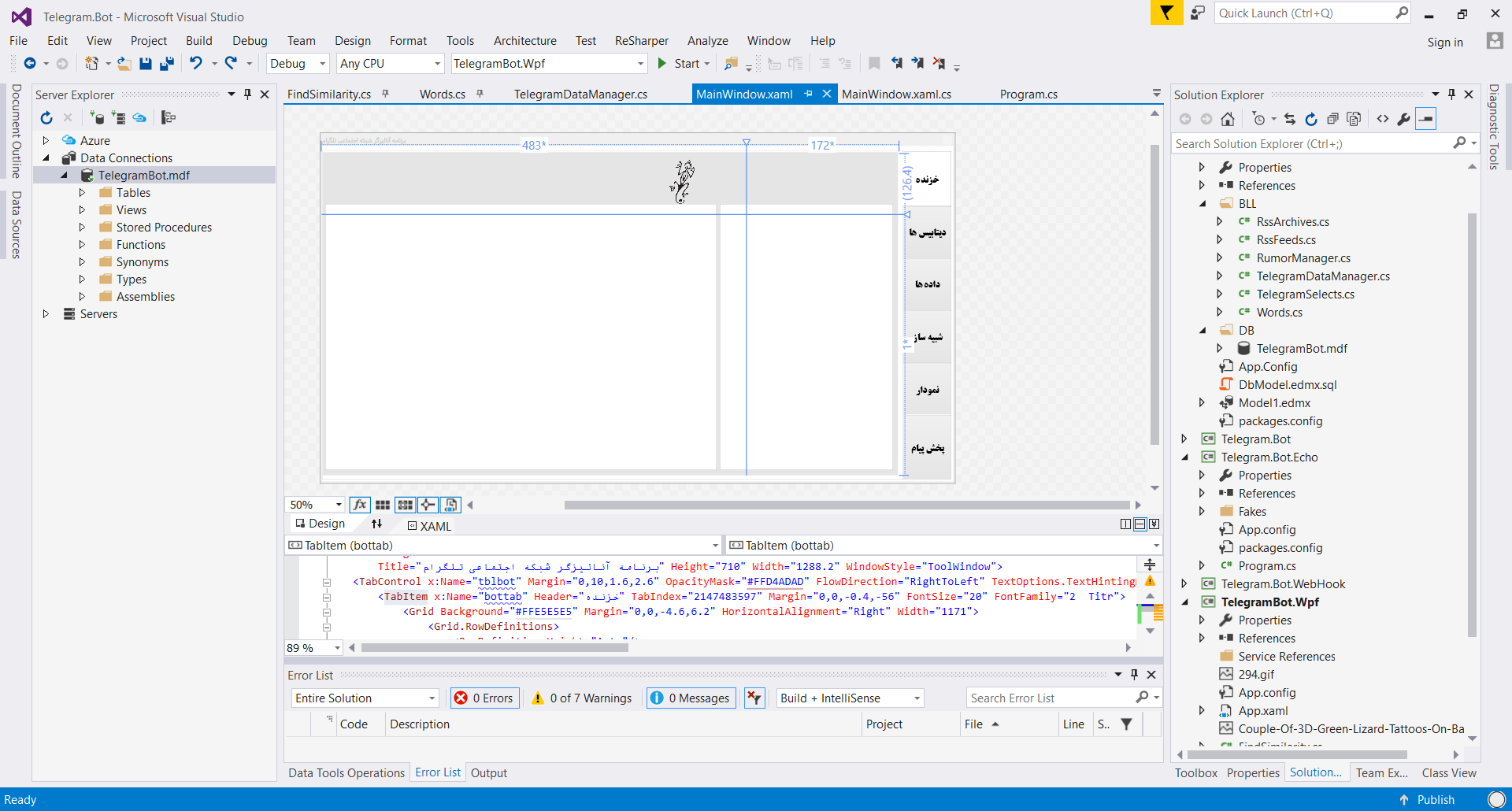
### بخش نمایش دیتابیس های موضوعی:



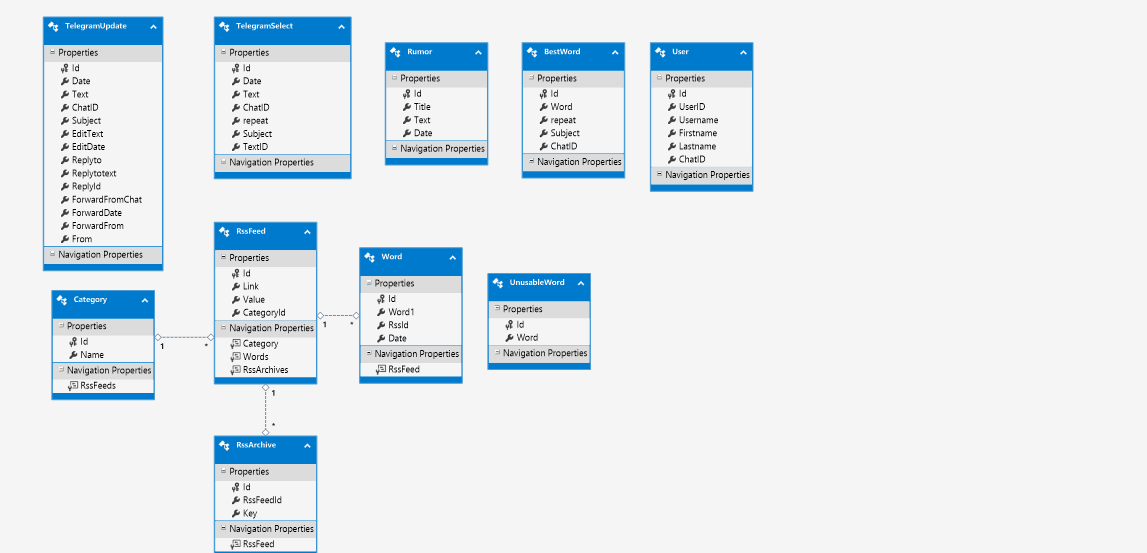
### صفحه ابتدایی برنامه مربوط به تمام لاگ های رباط در هنگام خزش و جدیدترین منابع آپدیت دیتابیس موضوعی



### تصویری از محیط پروژه در ویژوال استدیو 2015



### مودلینگ یا طرح کلی پشت پرده ی دیتابیس ها

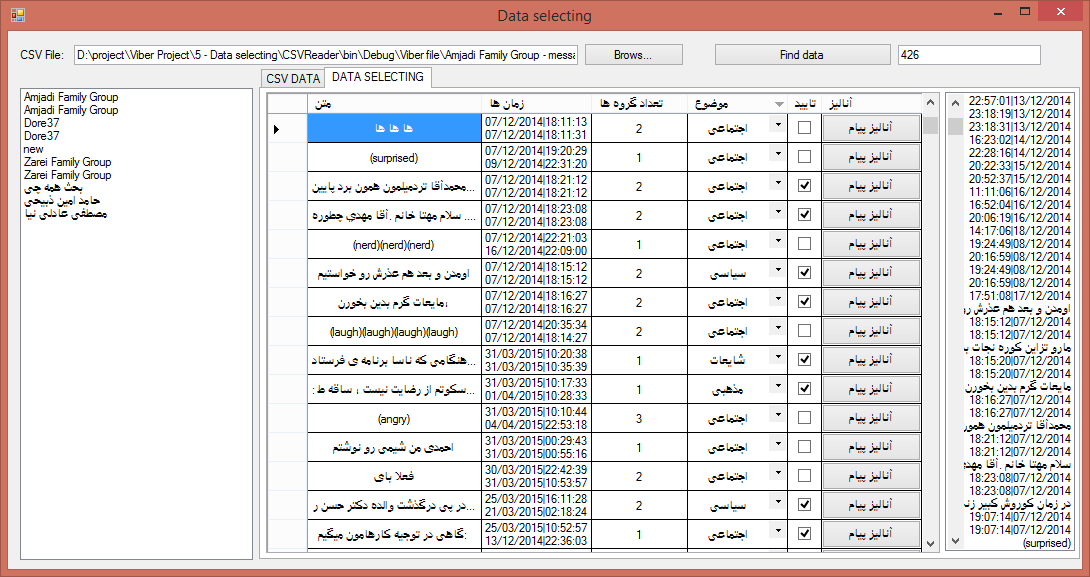


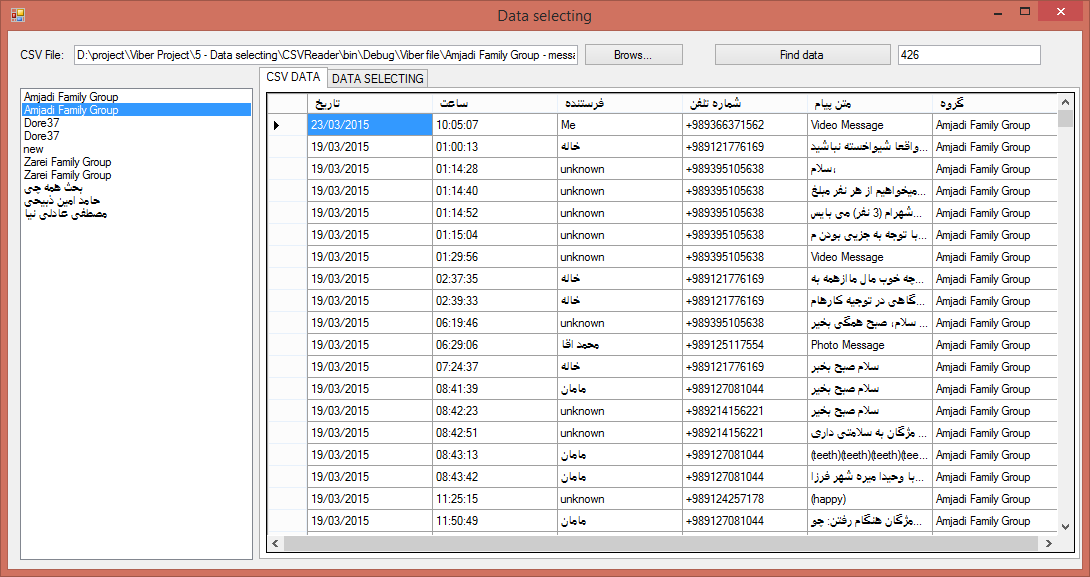
پشت پرده و مپینگ یا مدلینگ جداول دیتابیس در پروژه

شامل جداول:

1. تمام پیام ها با مشخصات خود و گروهشان
2. پیام های پر تکرار درسطح جامعه ی هدف به همراه مشخصاتشان
3. جداول دارای کلمات موضوعی گرفته شده از آر اس اس سایت های خبری برای موضوع دهی به پیام ها
4. جدول شایعات برای شناسایی پیام های شایعه
5. جدول موضوعات
6. جدول منبع آر اس اس ها برای دریافت اطلاعات
7. جدول کلمات هجو برای پالایش اطلاعات
8. جدول آرشیو آر اس اس های خوانده شده تا کنون
9. جدول تمام کاربران تلگرام در ارتباط با ربات به همراه اطلاعات شناسایی و گروهی که با ربات در آن هستند
10. جدول برترین کلمات از لحاظ استفاده شدن در پیام ها

### تصاویری از برنامه خواندن اطلاعات از روی فایل پشتیبانی وایبر





## مقاله های مربوط پیشینه ی تحقیق:

### Bigdata چیست و چرا اهمیت دارد؟

Big Data اصطلاحی رایج است که رشد و در دسترس بودن داده، چه ساختارمند و چه غیرساختارمند، را توصیف می ­کند. Big Data ممکن است به اندازه اینترنت برای کسب ­و­کار – و جامعه – مهم باشد. چرا؟ داده ­های بیشتر به تحلیل ­های دقیق ­تر می ­انجامد. تحلیل ­های دقیق ­تر منجر به تصمیم­ گیر ی­های مطمئن بیشتری شده و تصمیمات بهتر، می ­تواند معنای کارایی بیشتر عملیات، کاهش هزینه ­ها و کاهش ریسک­ ها باشد.  
  
  
Big Data تعریف می ­شود  
در سال 2001، تحلیلگر صنعت، داگ لنی (که اکنون با گارتنر کار می­ کند)، Big Data را به صورت سه V تعریف کرد: حجم (Volume)، سرعت (Velocity) و تنوع (Variety).

حجم: فاکتورهای بسیاری به افزایش حجم داده ­ها کمک می­ کند. داده­ های بر پایه تراکنش ذخیره شده در طی سالیان، داده ­های غیرساختارمند سرازیر شده از رسانه ­های اجتماعی؛ مقدار در حال افزایش داده­ های ماشین-به-ماشین و سنسور جمع ­آوری شده. در گذشته، حجم انبوه داده یک مسئله ذخیره کردن بود. اما با کاهش هزینه ­های ذخیره، مسائل دیگری سر بر می ­آورند؛ شامل چگونگی تعیین ارتباط در حجم زیاد داده­ ها و چگونگی استفاده از علم تجزیه و تحلیل به منظور ایجاد ارزش از داده ­های مرتبط.

سرعت: داده­ ها با سرعتی بی ­سابقه وارد شده و باید در زمان مناسب به سراغ آن­ها رفت. تگ­ های RFID، سنسورها و اندازه­ گیری هوشمند، نیاز به سر و کله زدن با جریانات داده را در اولین زمان نزدیک به اکنون را ایجاد می­ کنند. واکنش سریع به کار با سرعت داده ­ها، چالشی برای بیشتر سازمان­ هاست.

تنوع: داده­ ها به شکل­ های گوناگونی وارد می­ شوند. داده ­های عددی ساختاریافته در پایگاه­ های داده سنتی؛ اطلاعات ایجاد شده از برنامه­ های کاربردی کسب ­و­کار؛ اسناد متنی غیرساختاریافته، ایمیل، صدا و تراکنش­ های مالی. مدیریت، ادغام و حاکمیت بر انواع گوناگون داده، چیزی است که بسیاری از سازمان­ ها هنوز با آن درگیرند.

چرا Big Data باید برای شما مهم باشد؟  
مسئله واقعی این نیست که مقدار زیادی داده به دست آورید؛ این است که با آن چه می­ کنید. دیدگاه امیدوارانه این است که سازمان­ ها قادر به تحصیل داده از هر منبعی بوده، داده­ های مرتبط را تهیه کرده و آن را تحلیل کنند تا پاسخ سؤالاتی را بیابند که 1) کاهش هزینه­ ها، 2) کاهش زمان، 3) توسعه محصولات جدید و پیشنهادات جدید، و 4) تصمیم ­گیری هوشمندانه ­تر کسب ­وکار را مقدور می ­سازند. برای مثال، با ترکیب Big Data و تحلیل­ های قوی، این امکان وجود دارد تا:

علت های اصلی شکست ها، مسائل و نقوص را در لحظه تعیین کرد تا سالانه تا میلیاردها دلار صرفه­ جویی کرد.

مسیر وسیله ­های حمل بسته­ های تحویلی را زمانی که هنوز در جاده هستند، بهینه کرد.

در چند دقیقه تمام سبد ریسک را دوباره حساب کرد.

سریعاً مشتریانی که بیشترین اهمیت را دارند، شناسایی کرد.

......

Big Data واژه ­ای است برای مجموع ه­ای از ست داده ­های بسیار بزرگ و پیچیده، که استفاده از ابزارهای مدیریت پایگاه داده در دست و یا برنامه ­های کاربردی سنتی پردازش داده، برای پردازش آن­ها دشوار خواهد بود. چالش­ ها شامل استخراج، Curation، ذخیره ­سازی، جستجو، اشتراک، انتقال، آنالیز و بصری ­سازی است. در سال 2012، محدودیت اندازه ست داده ­ها، با زمان پردازش معقول، بر اگزابایت، میلیون ترابایت، قرار داشت.  
کار با Big Data با استفاده از سیستم ­های مدیریت دیتابیس ­های رابطه ­ای و بسته ­های بصری ­سازی و تحلیل ­های دسکتاپ، دشوار بوده و نیازمند نرم ­افزار بسیار موازی در حال کار بر روی ده­ ها، صدها و یا حتی هزاران سرور هستند. آنچه که Big Data شناخته می ­شود، بنا بر قابلیت ­های سازمان مدیریت کننده آن، و قابلیت­ های برنامه ­های کاربردی که به طور سنتی در آن زمینه داده پردازش و تحلیل می­ کنند، متفاوت است. برای برخی سازمان­ ها، رویارویی با صدها گیگابایت داده برای اولین بار ممکن است نیاز به بازبینی آپشن ­های مدیریت داده را ایجاد کند. برای برخی دیگر، ممکن است تا ده ­ها و صدها ترابایت طول بکشد که سایز داده به موضوعی قابل توجه تبدیل شود.

منبع : از انجمن سایت http://www.developercenter.ir/

### ربات تلگرام چیست؟

روبات تلگرام نوعی حساب کاربری ویژه است که برای ارسال و دریافت اتوماتیک پیام طراحی شده است. کاربران میتوانند با ارسال دستورات گوناگون با ربات های تلگرام در ارتباط باشند، این ارتباط میتواند در یک مکالمه شخصی و یا در یک گروه ایجاد شود. شما میتوانید ربات خود را با استفاده از API تلگرام و از طریق درخواست های HTTPS مدیریت کنید.

با توجه به افزایش امکانات ربات تلگرام و درخواست مکرر کاربران وبسیما برای ارائه فیلم های آموزشی، در حال تهیه و انتشار مجموعه آموزشی هستیم که هم اکنون در بخش[آموزش کدنویسی تلگرام](http://school.websima.com/downloads/category/%DA%A9%D8%AF%D9%86%D9%88%DB%8C%D8%B3%DB%8C-%D8%B1%D8%A8%D8%A7%D8%AA-%D8%AA%D9%84%DA%AF%D8%B1%D8%A7%D9%85/)در دسترس شما عزیزان قرار گرفته است.

آموزش های ارائه شده با جزئیات بالا تهیه شده و حتی برای افرادی که دانش بسیار کمی از کدنویسی دارند قابل استفاده است.

## تفاوت های ربات تلگرام با افراد

* ربات ها آخرین زمان آنلاین بودن ندارند و همیشه عبارت bot برای آنها نمایش داده میشود.
* ربات ها فضای محدودی برای ذخیر اطلاعات داردند و پیام های قدیمی تر پس از پردازش توسط شما حذف میشوند.
* ربات تلگرام نمیتواند شروع کننده مکالمه با کاربران باشد، تنها زمانیکه کاربر ربات را به گروه اضافه کند یا به آن پیام بفرستد امکان ارسال پیام از طرف ربات وجود خواهد داشت.
* در انتهای نام کاربری ربات همیشه کلمه bot حضور دارد.
* با اضافه شدن ربات به گروه، همه پیام ها توسط ربات دریافت و تحلیل نمیشود مگر آنکه تنظیمات آن را تغییر داده باشید.
* ربات های تلگرام هرگز غیرفعال، آفلاین یا خسته نمیشوند!

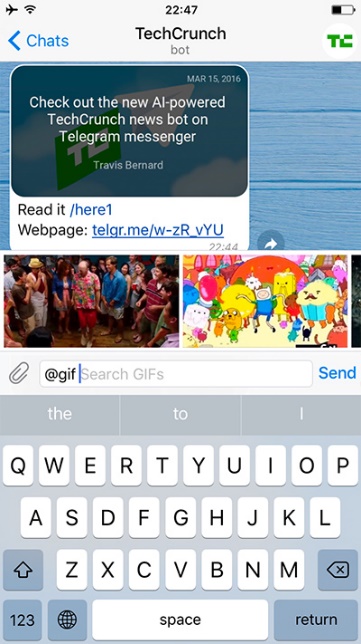
منبع : (ربات تلگرام بدون تاريخ)/

### Bots: An introduction for developers

(آموزش و معرفی رسمی خود سایت تلگرام درباره ربات های تلگرام)

Bots are third-party applications that run inside Telegram. Users can interact with bots by sending them messages, commands and [inline requests](https://core.telegram.org/bots#inline-mode). You control your bots using HTTPS requests to our [bot API](https://core.telegram.org/bots/api).

### 1. What can I do with bots?

[](https://core.telegram.org/file/811140081/1/VldnlW70g2I/a140b0216e7d156ebc)

To name just a few things, you could use bots to:

* **Get customized notifications and news**. A bot can act as a smart newspaper, sending you relevant content as soon as it's published.  
  [**Forbes Bot**](https://telegram.me/ForbesBot), [**TechCrunch Bot**](https://telegram.me/TechCrunchBot)
* **Integrate with other services**. A bot can enrich Telegram chats with content from external services.  
  [**Image Bot**](https://telegram.me/imagebot), [**GIF bot**](https://telegram.me/gif), [**IMDB bot**](https://telegram.me/imdb), [**Wiki bot**](https://telegram.me/wiki), [**Music bot**](https://telegram.me/music), [**Youtube bot**](https://telegram.me/youtube), [**GitHub bot**](https://telegram.me/githubbot)
* **Create custom tools**. A bot may provide you with alerts, weather forecasts, translations, formatting or other services.  
  [**Markdown bot**](https://telegram.me/Bold), [**Sticker bot**](https://telegram.me/sticker), [**Vote bot**](https://telegram.me/vote) **(NEW)**, [**Like bot**](https://telegram.me/like) **(NEW)**
* **Build single- and multiplayer games**. A bot can play chess and checkers against you, act as host in quiz games, or even take up the dungeon master's dice for an RPG.  
  [**Trivia bot**](https://telegram.me/TriviaBot)
* **Build social services**. A bot could connect people looking for conversation partners based on common interests or proximity.  
  [**HotOrBot**](https://telegram.me/hotorbot)
* **Do virtually anything else**. Except for dishes — bots are terrible at doing the dishes.

### 2. How do bots work?

At the core, Telegram Bots are special accounts that do not require an additional phone number to set up. Users can interact with bots in two ways:

* Send messages and [commands](https://core.telegram.org/bots#commands) to bots by opening a chat with them or by adding them to groups. This is useful for chat bots or news bots like the official [TechCrunch](https://telegram.me/techcrunchbot) and [Forbes](https://telegram.me/forbesbot) bots.
* Send requests directly from the input field by typing the bot's @username and a query. This allows sending content from [inline bots](https://core.telegram.org/bots/inline) directly into any chat, group or channel.

Messages, commands and requests sent by users are passed to the software running on your servers. Our intermediary server handles all encryption and communication with the Telegram API for you. You communicate with this server via a simple HTTPS-interface that offers a simplified version of the Telegram API. We call that interface our [Bot API](https://core.telegram.org/bots/api).

A detailed description of the Bot API is available on [this page »](https://core.telegram.org/bots/api)

### 3. How do I create a bot?

[](https://core.telegram.org/file/811140327/1/zlN4goPTupk/9ff2f2f01c4bd1b013)

There's a… bot for that. Just talk to [BotFather](https://telegram.me/botfather) (described [below](https://core.telegram.org/bots#botfather)) and follow a few simple steps. Once you've created a bot and received your authorization token, head down to the [Bot API manual](https://core.telegram.org/bots/api) to see what you can teach your bot to do.

You may also like to check out some **code examples** [here »](https://core.telegram.org/bots/samples)

### 4. How are bots different from humans?

* Bots have no online status and no last seen timestamps, the interface shows the label **‘bot’** instead.
* Bots have limited cloud storage — older messages may be removed by the server shortly after they have been processed.
* Bots can't initiate conversations with users. A user **must** either add them to a group or send them a message first. People can use telegram.me/<bot\_username> links or username search to find your bot.
* Bot usernames always end in ‘bot’ (e.g. [@TriviaBot](https://telegram.me/triviabot), [@GitHub\_bot](https://telegram.me/githubbot)).
* When added to a group, bots do not receive all messages by default (see [Privacy mode](https://core.telegram.org/bots#privacy-mode)).
* Bots never eat, sleep or complain (unless expressly programmed otherwise).

### 5. Bot perks

Telegram bots are unique in many ways — we offer [two](https://core.telegram.org/bots#keyboards) [kinds](https://core.telegram.org/bots#inline-keyboards-and-on-the-fly-updating) of keyboards, additional interfaces for [default commands](https://core.telegram.org/bots#global-commands) and [deep linking](https://core.telegram.org/bots#deep-linking) as well as [markdown support](https://core.telegram.org/bots#markdown-bold-italic-and-url-text) and much, much more.

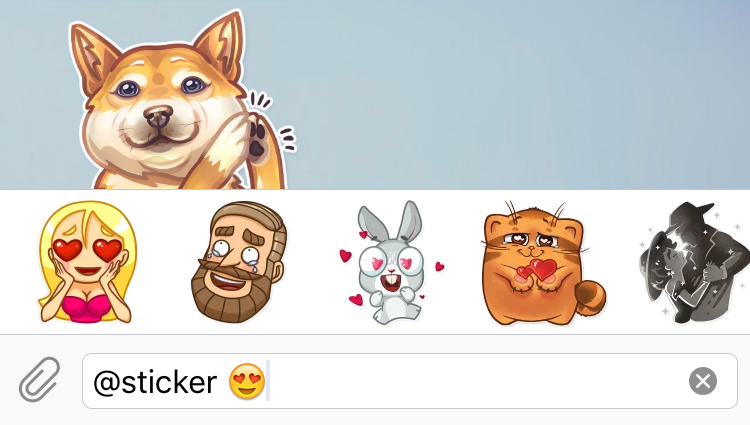
#### Inline mode

[](https://core.telegram.org/file/811140221/1/fW9vnLya4Fg/e2b5c530c7b0e019c4)

Users can interact with your bot via [**inline queries**](https://core.telegram.org/bots/api#inline-mode) straight from the **text input field** in**any** chat. All they need to do is start a message with your bot's username and then type a query.

Having received the query, your bot can return some results. As soon as the user taps one of them, it will be sent to the user's currently opened chat. This way, people can request content from your bot in any of their chats, groups or channels.

Check out this [blog](https://telegram.org/blog/inline-bots) to see a sample inline bot in action. You can also try the [@sticker](https://telegram.me/sticker)and [@music](https://telegram.me/music) bots to see for yourself.

[](https://core.telegram.org/file/811140558/1/POjp00-nHqE/50d0312845a05e6da9)

We've also implemented an easy way for your bot to [switch between inline and PM modes](https://core.telegram.org/bots/inline#switching-inline-pm-modes).

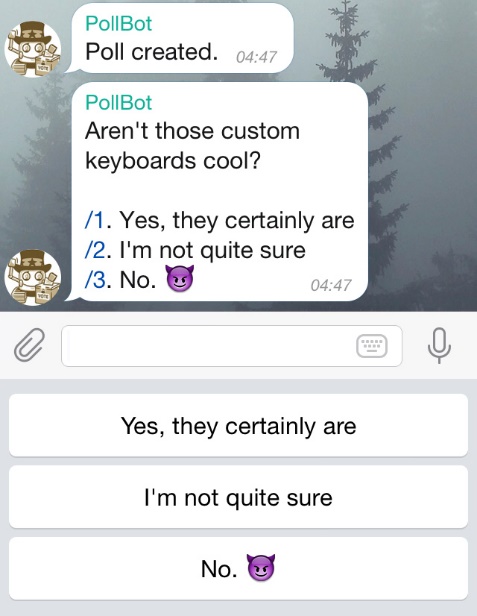
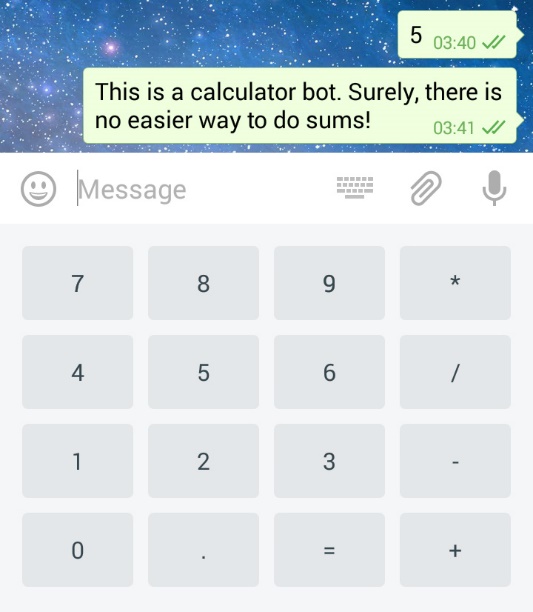
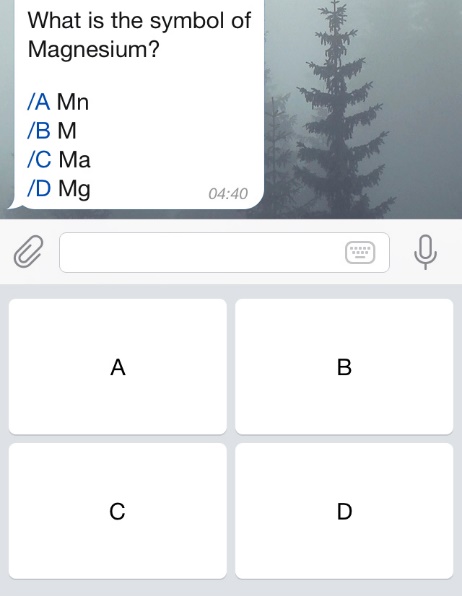
[Read more about the inline mode »](https://core.telegram.org/bots/inline)

#### Keyboards

Traditional chat bots can of course be taught to understand human language. But sometimes you want some more formal input from the user — and this is where **custom keyboards** can become extremely useful.

Whenever your bot sends a message, it can pass along a special keyboard with predefined reply options (see[ReplyKeyboardMarkup](https://core.telegram.org/bots/api/#replykeyboardmarkup)). Telegram apps that receive the message will display your keyboard to the user. Tapping any of the buttons will immediately send the respective command. This way you can drastically simplify user interaction with your bot.

We currently support text and emoji for your buttons. Here are some custom keyboard examples:

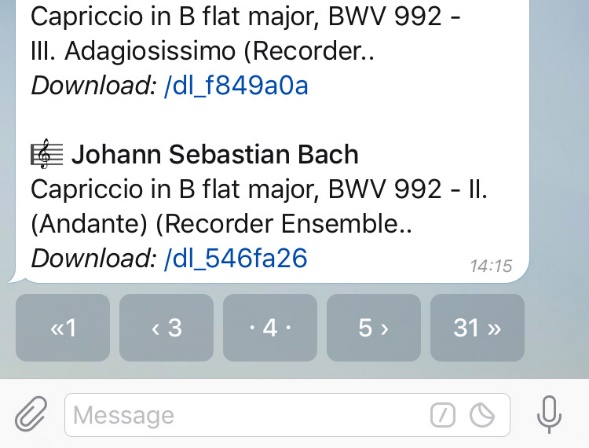
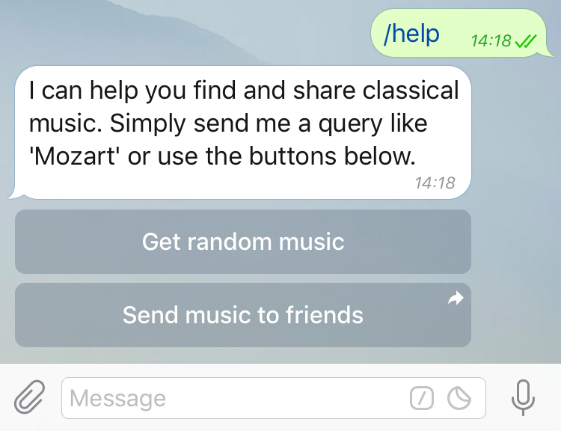
[](https://core.telegram.org/file/811140184/1/5YJxx-rostA/ad3f74094485fb97bd) [](https://core.telegram.org/file/811140880/1/jS-YSVkDCNQ/b397dfcefc6da0dc70) [](https://core.telegram.org/file/811140733/2/KoysqJKQ_kI/a1ee46a377796c3961)

For more technical information on custom keyboards, please consult the [Bot API manual](https://core.telegram.org/bots/api) (see [sendMessage](https://core.telegram.org/bots/api#sendmessage)).

#### Inline keyboards and on-the-fly updating

There are times when you'd prefer to do things without sending any messages to the chat. For example, when your user is changing settings or flipping through search results. In such cases you can use Inline Keyboards that are integrated directly into the messages they belong to.

Unlike with custom reply keyboards, pressing buttons on inline keyboards doesn't result in messages sent to the chat. Instead, inline keyboards support buttons that work behind the scenes: [callback buttons](https://core.telegram.org/bots/2-0-intro#callback-buttons), [URL buttons](https://core.telegram.org/bots/2-0-intro#url-buttons) and[switch to inline buttons](https://core.telegram.org/bots/2-0-intro#switch-to-inline-buttons).

[](https://core.telegram.org/file/811140217/1/NkRCCLeQZVc/17a804837802700ea4) [](https://core.telegram.org/file/811140659/1/RRJyulbtLBY/ea6163411c7eb4f4dc) [](https://core.telegram.org/file/811140999/1/2JSoUVlWKa0/4fad2e2743dc8eda04)

When callback buttons are used, your bot can update its existing messages (or just their keyboards) so that the chat remains tidy. Check out these sample bots to see inline keyboards in action: [@music](https://telegram.me/music), [@vote](https://telegram.me/vote), [@like](https://telegram.me/like).

[Read more about inline keyboards and on-the-fly editing »](https://core.telegram.org/bots/2-0-intro#new-inline-keyboards)

#### Commands

Commands present a more flexible way to communicate with your bot. The following syntax may be used:

/command [optional] [argument]

A command must always start with the ‘/’ symbol and may not be longer than 32 characters. Commands can use latin letters, numbers and underscores. Here are a few examples:

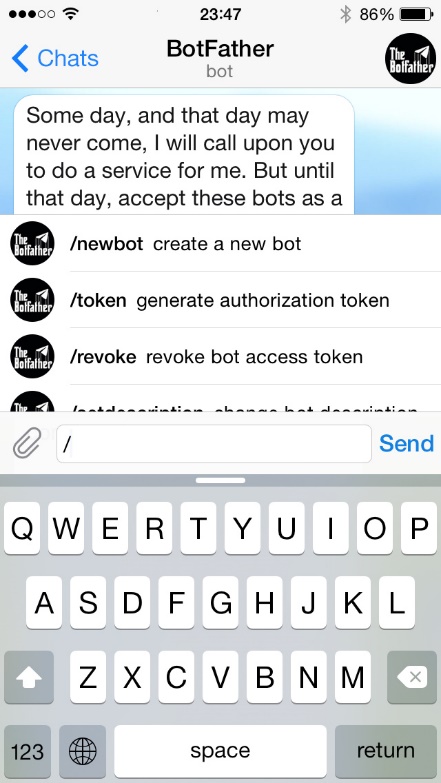
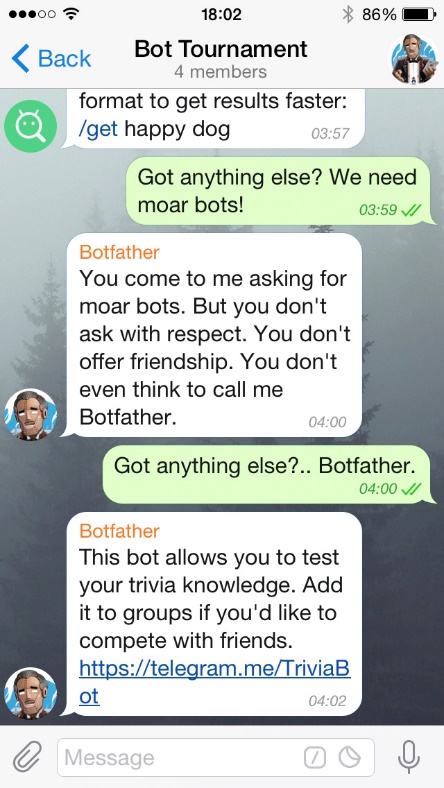
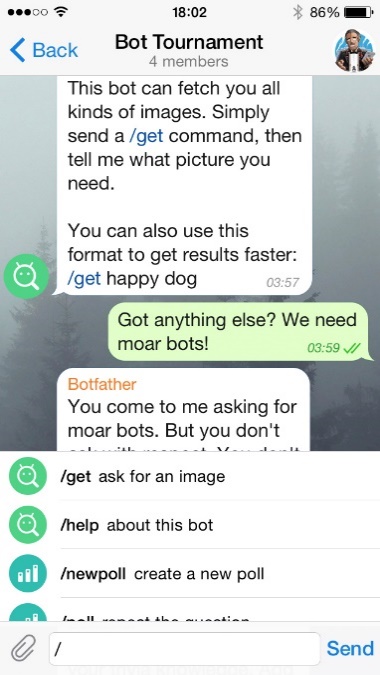
/get\_messages\_stats

/set\_timer 10min Alarm!

/get\_timezone London, UK

Messages that start with a slash will be always passed to the bot (along with replies to its messages and messages that @mention the bot by username). Telegram apps will:

* Suggest a list of [supported commands](https://core.telegram.org/bots#edit-settings) with descriptions when the user enters a ‘/’ (for this to work, you need to have provided a list of commands to the [BotFather](https://core.telegram.org/bots#botfather)). Tapping on a command in the list immediately sends the command.
* Show an additional **(/)** button in the input field in all chats with bots. Tapping it types a ‘/’ and shows the list of commands.
* Highlight **/commands** in messages. When the user taps a highlighted command, the command is sent at once.

[](https://core.telegram.org/file/811140845/2/rNUxpcGDeQU/05eaaf20b0dbaf9cb3) [](https://core.telegram.org/file/811140315/2/gf7_D2HbeyM/e3ca2de4de7918f826) [](https://core.telegram.org/file/811140029/1/s5zv4fbWdhw/a04aefa0ee0557f16a) 

If multiple bots are in a group, it is possible to add bot usernames to commands in order to avoid confusion:

/start@TriviaBot

/start@ApocalypseBot

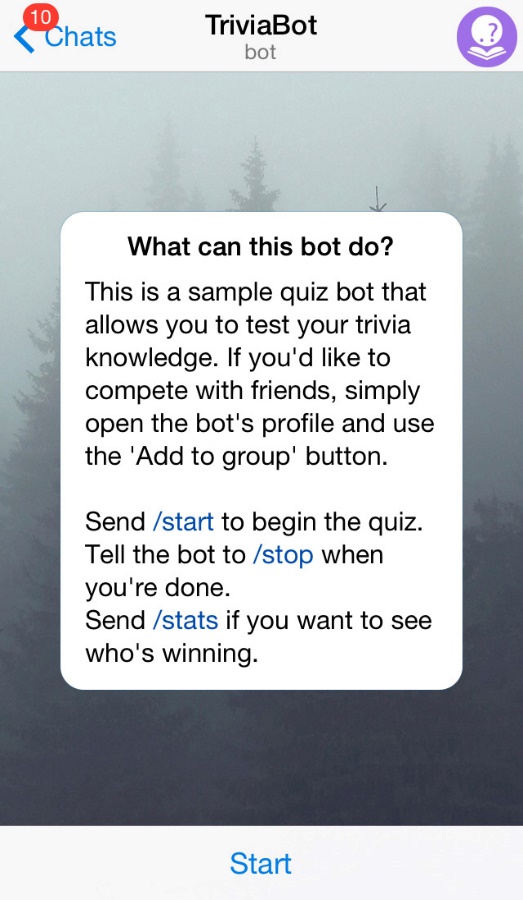
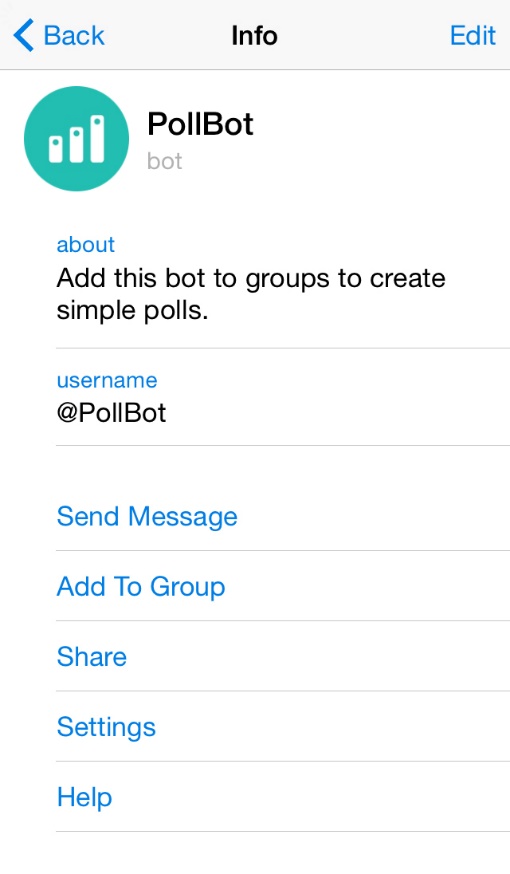
This is done automatically when commands are selected via the list of suggestions. Please remember that your bot needs to be able to process commands that are followed by its username.

##### Global commands

In order to make it easier for users to navigate the bot multiverse, we ask all developers to support a few basic commands. Telegram apps will have **interface shortcuts** for these commands.

* **/start** - begins interaction with the user, e.g., by sending a greeting message. This command can also be used to pass additional parameters to the bot (see [Deep linking](https://core.telegram.org/bots#deep-linking))
* **/help** - returns a help message. It can be a short text about what your bot can do and a list of commands.
* **/settings** - (if applicable) returns the bot's settings for this user and suggests commands to edit these settings.

Users will see a **Start** button when they first open a conversation with your bot. **Help** and **Settings** links will be available in the menu on the bot's profile page.

[](https://core.telegram.org/file/811140979/2/yD8AphHbahk/7662d14f4e0442ae3a) [](https://core.telegram.org/file/811140479/2/1c2zUWhR7sA/98889b2a45f8e42a35)

#### Formatting: bold, italic, fixed-width text and inline links

You can use bold, italic or fixed-width text, as well as inline links in your bots' messages. Telegram clients will render them accordingly.

[Read more in the Bot API manual »](https://core.telegram.org/bots/api#formatting-options)

#### Privacy mode

Bots are frequently added to groups in order to augment communication between human users, e.g. by providing news, notifications from external services or additional search functionality. This is especially true for work-related groups. Now, when you share a group with a bot, you tend to ask yourself “How can I be sure that the little rascal isn't selling my chat history to my competitors?” The answer is — **privacy mode**.

A bot running in privacy mode will not receive all messages that people send to the group. Instead, it will only receive:

* Messages that start with a slash ‘/’ (see [Commands](https://core.telegram.org/bots#commands) above)
* Replies to the bot's own messages
* Service messages (people added or removed from the group, etc.)

On one hand, this helps some of us sleep better at night (in our tinfoil nightcaps), on the other — it allows the bot developer to save a lot of resources, since they won't need to process tens of thousands irrelevant messages each day.

Privacy mode is enabled by default for all bots, except bots that were added to the group as **admins**. It can be disabled, so that the bot will begin receiving all messages like an ordinary user. We only recommend doing this in cases where it is absolutely necessary for your bot to work — users can always see a bot's current privacy setting in the group members list. In most cases, using the [force reply](https://core.telegram.org/bots/api#forcereply) option for the bot's messages should be more than enough.

[So what messages exactly will my bot get? »](https://core.telegram.org/bots/faq#what-messages-will-my-bot-get)

#### Deep linking

Telegram bots have a [deep linking](https://en.wikipedia.org/wiki/Deep_linking) mechanism, that allows for passing additional parameters to the bot on startup. It could be a command that launches the bot — or an auth token to connect the user's Telegram account to their account on some external service.

Each bot has a link that opens a conversation with it in Telegram — https://telegram.me/<bot username>. You can add the parameters **start** or **startgroup** to this link, with values up to 64 characters long. For example:

https://telegram.me/triviabot?startgroup=test

A-Z, a-z, 0-9, \_ and - are allowed. We recommend using [base64url](https://en.wikipedia.org/wiki/Base64#URL_applications) to encode parameters with binary and other types of content.

Following a link with the **start** parameter will open a one-on-one conversation with the bot, showing a START button in the place of the input field. If the **startgroup** parameter is used, the user will be prompted to select a group to add the bot to. As soon as a user confirms the action (presses the START button in their app or selects a group to add the bot to), your bot will receive a message from that user in this format:

/start PAYLOAD

PAYLOAD stands for the value of the **start** or **startgroup** parameter that was passed in the link.

##### Deep linking Example

Suppose the website example.com would like to send notifications to its users via a Telegram bot. Here's what they could do to enable notifications for a user with the ID 123.

1. [Create a bot](https://core.telegram.org/bots#botfather) with a suitable username, e.g. @ExampleComBot
2. Set up a [webhook](https://core.telegram.org/bots/api#setwebhook) for incoming messages
3. Generate a random string of a sufficient length, e.g. $memcache\_key = "vCH1vGWJxfSeofSAs0K5PA"
4. Put the value 123 with the key $memcache\_key into Memcache for 3600 seconds (one hour)
5. Show our user the button https://telegram.me/ExampleComBot?start=vCH1vGWJxfSeofSAs0K5PA
6. Configure the webhook processor to query Memcached with the parameter that is passed in incoming messages beginning with /start. If the key exists, record the chat\_id passed to the webhook as**telegram\_chat\_id** for the user 123. Remove the key from Memcache.
7. Now when we want to send a notification to the user 123, check if they have the field **telegram\_chat\_id**. If yes, use the [sendMessage](https://core.telegram.org/bots/api#sendmessage) method in the [Bot API](https://core.telegram.org/bots/api) to send them a message in Telegram.

#### Location and Number

Some bots need extra data from the user to work properly. For example, knowing the user‘s location helps provide more relevant geo-specific results. The user’s phone number can be very useful for integrations with other services, like banks, etc.

Bots can ask a user for their **location** and **phone number** using special buttons. Note that both phone number and location request buttons will only work in private chats.

[](https://core.telegram.org/file/811140587/2/jaowDLZg2l0/5ba3f7d7fd5c6c28dc)

When these buttons are pressed, Telegram clients will display a confirmation alert that tells the user what's about to happen.

[Manual: Number and location buttons »](https://core.telegram.org/bots/api#keyboardbutton)

### 6. BotFather

Jump to top to learn everything about [Telegram bots »](https://core.telegram.org/bots)

[BotFather](https://telegram.me/botfather) is the one bot to rule them all. It will help you create new bots and change settings for existing ones.

#### Create a new bot

Use the **/newbot** command to create a new bot. The BotFather will ask you for a name and username, then generate an authorization token for your new bot.

The **name** of your bot will be displayed in contact details and elsewhere.

The **Username** is a short name, to be used in mentions and telegram.me links. Usernames are 5-32 characters long and are case insensitive, but may only include Latin characters, numbers, and underscores. Your bot's username**must** end in ‘bot’, e.g. ‘tetris\_bot’ or ‘TetrisBot’.

The **token** is a string along the lines of 110201543:AAHdqTcvCH1vGWJxfSeofSAs0K5PALDsaw that will be required to authorize the bot and send requests to the [Bot API](https://core.telegram.org/bots/api).

#### Generate an authorization token for your bot

If your existing token is compromised or you lost it for some reason, use the **/token** command to generate a new one.

#### Edit settings

The remaining commands are pretty self-explanatory:

* **/setname** – change your bot's name.
* **/setdescription** — changes the bot's **description**, a short text of up to 512 characters, describing your bot. Users will see this text at the beginning of the conversation with the bot, titled ‘What can this bot do?’.
* **/setabouttext** — changes the bot's **about info**, an even shorter text of up to 120 characters. Users will see this text on the bot's profile page. When they share your bot with someone, this text will be sent together with the link.
* **/setuserpic** — changes the bot‘s profile pictures. It’s always nice to put a face to a name.
* **/setcommands** — changes the list of commands supported by your bot. Each command has a name (must start with a slash ‘/’, alphanumeric plus underscores, no more than 32 characters, case-insensitive), parameters, and a text description. Users will see the list of commands whenever they type ‘/’ in a conversation with your bot.
* **/setjoingroups** — determines whether your bot can be added to groups or not. Any bot must be able to process private messages, but if your bot was not designed to work in groups, you can disable this.
* **/setprivacy** — determines which messages your bot will receive when added to a group. With privacy mode disabled, the bot will receive all messages. We recommend leaving [privacy mode](https://core.telegram.org/bots#privacy-mode) enabled.
* **/deletebot** — deletes your bot and frees its username.

Please note, that it may take **a few minutes** for changes to take effect.

منبع : سایت رسمی تلگرام: <https://core.telegram.org/bots>

توضیحاتی درباره ی Api[[15]](#footnote-15) های تلگرام از سایت رسمی همان شرکت:

### Telegram APIs

We offer two kinds of APIs for developers. The [Bot API](https://core.telegram.org/#bot-api) allows you to easily create programs that use Telegram messages for an interface. The [Telegram API](https://core.telegram.org/#telegram-api) allows you to build your own customized Telegram clients. You are welcome to use both APIs free of charge.

### Bot API

This API allows you to connect bots to our system. Telegram Bots are special accounts that do not require an additional phone number to set up. These accounts serve as an interface for code running somewhere on your server.

To use this, you don't need to know anything about how our MTProto encryption protocol works — our intermediary server will handle all encryption and communication with the Telegram API for you. You communicate with this server via a simple HTTPS-interface that offers a simplified version of the Telegram API.

[**Learn more about the Bot API here »**](https://core.telegram.org/bots)

### Telegram API

This API allows you to build your own customized Telegram clients. It is 100% open for all developers who wish to create Telegram applications on our platform. Feel free to study the open [source code](https://telegram.org/apps#source-code) of existing Telegram applications for examples of how things work here. Don't forget to [register](https://core.telegram.org/api/obtaining_api_id) your application in our system.

* [Getting Started](https://core.telegram.org/#getting-started)
* [Security](https://core.telegram.org/#security)
* [Optimization](https://core.telegram.org/#optimization)
* [API methods](https://core.telegram.org/#api-methods)
* [Updates log](https://core.telegram.org/#updates-log)

منبع : https://core.telegram.org/

مقاله ای درباره EntityFramework که از آن در پروژه بهره برده شده است:

#### **Entity Framework چیست؟**

**Entity Framework** یک ORM است که توسط شرکت [Microsoft](http://www.microsoft.com/)توسعه داده میشود. نسخه 1.0 آن مایوس کننده بود و در موارد بسیاری کمبود داشت. اما با توجه به سرمایه گذاری زیادی که Microsoft روی پروژه انجام داده بود، مشخص بود که قصد جدی در توسعه این ابزار دارد. نسخه کنونی آن، 6 هست که با **2015 Visual Studio** روی سیستم نصب میشود. با توجه به اهمیت این بحث برای Microsoft، انتظار میرود که توسعه این ابزار به طور فعالی ادامه پیدا کند و امکانات کاملی به آن اضافه شود.

**نکته :**.**ORM** یا Object Relational Mapping ابزاری است که به وسیله آن می توان [جداول پایگاه داده](http://www.papro-co.ir/) را پردازش کرد و آنها را به فرمت شیءنگاشت کرد که در کدنویسی خواناتر باشند. در کدنویسی به جای این که به صورت مستقیم با جداول پایگاه داده کار کنیم بر روی اشیاء کار می کنیم. به این اشیاء **موجودیت یا Entity**گفته می شود.

ADO.NET Entity Framework به گونه ای طراحی شده است تا توسعه دهندگان را قادر سازد در مقابل مدل های مرسوم مفهومی (conceptual)، از یک مدل جدید با امکان دسترسی و برنامه نویسی مستقیم بر اساس ساختار یک پایگاه داده رابطه ای استفاده نمایند. [هدف از تولید](http://www.infobigbang.com/) این مدل، کاهش چشمگیر میزان کدنویسی، سهولت گسترش، ترمیم، بهبود و نگهداری برنامه ها بوده است.

**چرا از Entity Framework استفاده می کنیم؟**

·       [دسترسی به داده ها](http://www.papro-co.ir/) را ساده می سازد و به ما اجازه می دهد تا مدل مفهومی مورد نیاز برای کسب و کار را ایجاد کنیم.

·        Entity Framework اجازه میدهد تا شما روابط چندبه چند (n: n) را بدون join ها متعدد مدیریت نمایید.

·         جهت کاهش پیچیدگی کد، جداول مرتبط را در یک Entity مدل می کند.

·        پشتیبانی غنی و خوب از ارث بری ها برای جدول به ازای سلسله مراتب، زیرکلاس (subclass) و تایپ های به هم چسبیده (concrete type).

·        چون دارای مکانیسم بسیار قوی Query گیری از Entity هاست. کسانی که مایل با استفاده از LINQ نیستند، میتواننداز Entity Client provider  همراه باEntitySQLاستفاده نمایند.

·        مستقل از نوع پایگاه داده می باشد و بر اساس [استاندارد](http://www.fasana.ir/)ADO.NET Providers. کار می کند.

روش های Entity Framework

#### 1) روش Database First  :

یعنی وظیفه توسعه دهنده نرم افزار این است که پایگاه داده را ایجاد کند. سپس با دادناین [پایگاه داده به Entity Framework](http://www.papro-co.ir/)،  یک فایل مدل با پسوند edmx  تحویل بگیرد. در این فایل که یک فایل xmlاست، اطلاعات مربوط به موجودیتها، خصوصیات هر موجودیت و نوع داده هر خصوصیت قرار گرفته است. (به این قسمت از فایل مدل مدل مفهومی گفته می شود). در فایل مدل علاوه بر مدل مفهومی، مدل فیزیکی هم قرار دارد. مدل فیزیکی اطلاعات مربوط به جداول پایگاه داده و فیلدهای هر جدول و نیز نوع داده هر فیلد قرار گرفته است. فایل مدل شامل قسمت دیگری است که بیان می کند جداول پایگاه داده به چه موجودیتهایی نگاشت شده اند.

#### 2) روش Model First :

در این روش وظیفه توسعه دهنده نرم افزار، ایجاد مدل می باشد. یعنی او خود فایلedmxرا می سازد. بعد از این کار می تواند ازEFبخواهد تا پایگاه داده مربوط به این مدل را ایجاد کند. برای این کار او ازیک ابزار ویژوال به اسم Entity Data Model Designerاستفاده می کند. این ابزار ویژوال بسیار قدرتمند می باشدو امکاناتی از قبیل ساختن موجودیتها، تعیین ارتباط بین موجودیتها، ایجاد کلید برای هر موجودیت، کاردینالیتی رابطه ها و غیره را در اختیار توسعه دهنده قرار می دهد.

در روش های الف و ب ، در Code Behindمدل کلاسهایی مربوط به هر موجودیت و نیز [شی Context](http://www.papro-co.ir/) وجود دارد.

#### 3) روش Code First :

این روش خود شامل دو قسمت می باشد:

اول؛ این است که توسعه دهنده کلاس های مربوط به هر موجودیت را ایجاد می کند. او این کار را با برنامه نویسی انجام می دهد. یعنی باید تک تک کلاسها، به همراه خصوصیات هر کلاس و البته خصوصیاتNavigationرا بابرنامه نویسی ایجاد کند. خصوصیاتNavigationآن دسته از خصوصیاتی هستند که ارتباط آن کلاس را با کلاس های دیگر مشخص می کنند. در این روش هیچ فایل مدلی وجود نخواهد داشت.

دوم؛ این است که مثل روشDatabase Firstابتدا توسعه دهنده یک پایگاه داده ایجاد می کند، سپس مدل را ایجاد می کند، سپس از یک افزونه می خواهد که موجودیتها را به همراه Contextایجاد کند و اینها را از مدل جدا کند. اگر موجودیتها به این طریق ایجاد شوند به آنها POCOگفته می شود. فرق آن با روشDatabase Firstدر همین نکته است. در اینجا فایل مدل (فایلedmx ) فقط اطلاعات مربوط به موجودیتها را نگهداری می کند و در Code Behindآن هیچ کدی نوشته نمی شود.

**مزایا**  
1- سطح یادگیری ساده: کار با EF به نسبت بقیه ORMها تقریبا ساده تراست.  
2- پشتیبانی از پایگاه داده های مختلف در نسخه 4.  
3- توسعه فعال: همانطور که اشاره شد، به نظر می رسدMicrosoft، قصد جدی در توسعه این قابلیت دارد. اضافه شدن امکانات زیاد به نسخه 4 نشاندهنده همین مطلب هست. مثل امکان Code First، Lazy Loading و …که در نسخه 4 به EF اضافه شده اند.  
4-  توسعه دهندگان قادرند تا با ساختار شیء گرایانه سازگار با انواع سیستم های مدیریت داده کار کنند.  
5- پشتیبانی کامل از Linq.  
7-  مرتبط شدن چندین مدل مفهومی به یک ساختارپایگاه داده.

8-  برنامه های کاربردی قادر خواهند بود تا در مدل مفهومی خود (conceptual model) محوریت بیشتری به خود برنامه ببخشند به گونه ای که شیء گرایی، ارث بری، گونه ها (Type) های پیچیده تر و روابط بین جداول و اشیاء بهتر پیاده سازی شوند.  
9-برنامه نویس ها از کدنویسی زیادی پیچیده رها می شوند.  
10-  بدون کوچک ترین تغییر کد برنامه خواهیم توانست مدل ارتباط بین ساختار پایگاه داده و مدل مفهومی را تغییر دهیم.

**انتخاب بین Linq To SQL و EF:**

حال که با روش EF آشنا شدیم، در پایان بررسی می کنیم که چه زمانی از Linq و چه زمانی از EF استفاده کنیم.

الف) Linq to SQL

زمانی که:

1- از یک  ORMاستفاده کنیم و پایگاه داده ما در یک رابطه 1 به 1 با مدل اشیاء قرار دارد.

2- از یک  ORMکه ساختا رسلسله مراتبی ای را که دریک جدول ذخیره شده است، ارث بری کرده باشد.

3-استفاده از کلاس هایی که خودمان نوشتیم، درکنارکلاسهای تولید شده توسط سیستم یا ارثبری از کلاسهای پایه یا پیاده سازی رابط ها درکلاسها.

4- استفاده از  LINQبه عنوان ابزار  Queryنویسی

5- استفاده از یک  ORMبا حفظ کارآیی و امکان بهبود آن از طریقStored Procedureها وQueryهای کامپایل شده.

ب) Entity Framework

زمانی که :

1- بخواهیم برنامه ما به پایگاه داده های دیگری به جز  Microsoft SQL Serverنیز دسترسی پیدا کند.

2- بخواهیم  domain modelsبرای برنامه مان تعریف کنیم و از آن به عنوان لایه  persistenceاستفاده نماییم.

3- از یک راهکار  ORMی استفاده کنیم که ممکن است کلاسهایمان رابط های 1:1 با پایگاه داده داشته باشد یا ممکن است ساختاری بسیار متفاوت و پیچیده در ارتباط با پایگاه داده دارا باشد.

4- میخواهیم از یک  ORMاستفاده کنیم که ساختارهای سلسله مراتبی در آن، معادل های پایگاه داده ای متفاوتی داشته باشد. مانند وقتی که جدولی جهت نگهداری سلسله مراتب، جداولی برای هر یک ازکلاسها وجدولی جهت نگهداری گونه ای (Type) خاص داریم.

5- وقتی میخواهیم از  LINQجهت نوشتن Queryها استفاده کنیم درحالیکه  Queryدرپایگاه داده به شیوه خاص و درسلطه سازنده آن عمل میکند.

6- استفاده از کلاسهایی که خودمان نوشتیم، درکنارکلاسهای تولیدشده توسط سیستم یا ارثبری ازکلاسهای پایه یا پیادهسازی رابطها درکلاسها.

7- استفاده از یک  ORMبا حفظ کارآیی و امکان بهبود آن از طریقStored Procedureها و Queryهای کامپایل شده.

منبع : <http://papro-co.ir/entity-framework-%DA%86%DB%8C%D8%B3%D8%AA%D8%9F/>

درباره الگوریتم هایی برای رشته ها:

### Delta encoding

From Wikipedia, the free encyclopedia

*Not to be confused with*[*Elias delta coding*](https://en.wikipedia.org/wiki/Elias_delta_coding)*or*[*delta modulation*](https://en.wikipedia.org/wiki/Delta_modulation)*.*

**Delta encoding** is a way of storing or transmitting [data](https://en.wikipedia.org/wiki/Data) in the form of [*differences*](https://en.wikipedia.org/wiki/Data_differencing) (deltas) between sequential data rather than complete files; more generally this is known as [data differencing](https://en.wikipedia.org/wiki/Data_differencing). Delta encoding is sometimes called **delta compression**, particularly where archival histories of changes are required (e.g., in [revision control software](https://en.wikipedia.org/wiki/Revision_control_software)).

The differences are recorded in discrete files called "deltas" or "diffs". In situations where differences are small – for example, the change of a few words in a large document or the change of a few records in a large table – delta encoding greatly reduces data redundancy. Collections of unique deltas are substantially more space-efficient than their non-encoded equivalents.

From a logical point of view the difference between two data values is the information required to obtain one value from the other – see [relative entropy](https://en.wikipedia.org/wiki/Relative_entropy). The difference between identical values (under some [equivalence](https://en.wikipedia.org/wiki/Equivalence_relation)) is often called *0* or the [neutral element](https://en.wikipedia.org/wiki/Neutral_element).

### Levenshtein distance

From Wikipedia, the free encyclopedia

In [information theory](https://en.wikipedia.org/wiki/Information_theory) and [computer science](https://en.wikipedia.org/wiki/Computer_science), the **Levenshtein distance** is a [string metric](https://en.wikipedia.org/wiki/String_metric) for measuring the difference between two sequences. Informally, the Levenshtein distance between two words is the minimum number of single-character edits (i.e. insertions, deletions or substitutions) required to change one word into the other. It is named after [Vladimir Levenshtein](https://en.wikipedia.org/wiki/Vladimir_Levenshtein), who considered this distance in 1965.[[1]](https://en.wikipedia.org/wiki/Levenshtein_distance#cite_note-1)

Levenshtein distance may also be referred to as **edit distance**, although that may also denote a larger [family of distance metrics](https://en.wikipedia.org/wiki/Edit_distance).[[2]](https://en.wikipedia.org/wiki/Levenshtein_distance#cite_note-navarro-2):32 It is closely related to [pairwise string alignments](https://en.wikipedia.org/wiki/Sequence_alignment#Pairwise_alignment).

# Edit distance

From Wikipedia, the free encyclopedia

In [computer science](https://en.wikipedia.org/wiki/Computer_science), **edit distance** is a way of quantifying how dissimilar two [strings](https://en.wikipedia.org/wiki/String_(computing)) (e.g., words) are to one another by counting the minimum number of operations required to transform one string into the other. Edit distances find applications in [natural language processing](https://en.wikipedia.org/wiki/Natural_language_processing), where automatic [spelling correction](https://en.wikipedia.org/wiki/Spell_checker) can determine candidate corrections for a misspelled word by selecting words from a dictionary that have a low distance to the word in question. In [bioinformatics](https://en.wikipedia.org/wiki/Bioinformatics), it can be used to quantify the similarity of [DNA](https://en.wikipedia.org/wiki/DNA)sequences, which can be viewed as strings of the letters A, C, G and T.

Different definitions of an edit distance use different sets of string operations. The [Levenshtein distance](https://en.wikipedia.org/wiki/Levenshtein_distance) operations are the removal, insertion, or substitution of a character in the string. Being the most common metric, the Levenshtein distance is usually what is meant by "edit distance"

منبع : ویکی پدیا : <https://en.wikipedia.org/wiki/String-to-string_correction_problem>

### RSS چیست ؟

(خبر در نوك انگشتان شما)

RSS به قولی سرنام REALLY SIMPLE SYNDICATION و به قول كمی حرفه‌ای‌ترها سرنام RDF SITE SUMMARY است. ولی RSS سرنام هر چه باشد، پدیده‌ای است كه امكان جمع‌آوری اطلاعات و اخبار را از سایت‌های مختلف فراهم می‌كند بدون این كه مجبور باشید از این سایت به آن سایت سر بزنید. ویژگی RSS هم مثل بقیه تكنولوژی‌های خوب، در سادگی آن است :یک فرمت سریع و استاندارد، شبیه همان بولتن‌های خبری است منتها به شكل الكترونیك. امروزه اكثر سایت‌های وب مهم به این بولتن‌ها مجهز هستند، از وبلاگ‌های شخصی گرفته تا سایت‌های خبرگزاری‌های بزرگ دنیا. برای گرفتن این بولتن‌ها كافی است نرم‌افزار مناسب این كار را در اختیار داشته باشید و بولتن مورد نظر را مشترك شوید. شاید فكر كنید خواندن خبرها كه كاری ندارد، به صفحه اصلی سایت می‌رویم و خبرها را می‌خوانیم. اما حالتی را در نظر بگیرید كه بخواهید 50 سایت مختلف را ببینید و تازه معلوم نیست این سایت‌ها اصلاً خبر جدیدی داشته باشند یا نه. با RSS نیازی نیست به 50 تا سایت سر بزنید. كامپیوتر شما به طور خودكار و با زمان‌بندی مشخص با این سایت‌ها تماس می‌گیرد و جدیدترین خبرها و اطلاعات را دانلود می‌كند. بنابراین، فقط از یك پنجره می‌توانید تمام اتفاقات روز را مشاهده كنید. بدین ترتیب هم در وقت‌تان صرفه‌جویی كرده‌اید، و هم در این همه اطلاعات اضافی كه در وب وجود دارد گرفتار نشده‌اید.

**انتشار و جمع‌آوری**

RSS به دو شكل ایفای نقش می‌كند: وقتی كه خودتان اطلاعاتی دارید كه می‌خواهید منتشر كنید و در اختیار دیگران قرار دهید، كار <نشر> یا SYNDICATION را انجام می‌دهد و زمانی كه بخواهید اطلاعاتی را به دست آورده و مورد استفاده قرار دهید، كار <گردآوری> یا AGGREGATION را برعهده می‌گیرد. ابتدا به نقش گردآوری RSS می‌پردازیم و بعد توضیح می‌دهیم كه چگونه بولتنی خبری درست كنید كه دیگران بتوانند بخوانند.هیچ‌كس نیاز ندارد هر روز تمام خبرهای دنیا را بخواند، بلكه هر كسی بسته به علاقه، شغل، یا وظیفه‌ای كه دارد فقط بعضی از خبرها را می‌خواند. برای این منظور به یك ابزار گردآوری احتیاج داریم كه بتواند به طور خودكار به سراغ سایت‌هایی برود كه به اشتراك آنها درآمده‌ایم.ابزارهای گردآوری در بازار و اینترنت فراوان یافت می‌شود، هم رایگان و هم غیررایگان، هم مستقل و هم الحاقی (از جمله NEWSGATOR كه برای 2003OUTLOOK طراحی شده است). یكی از این برنامه‌هاFEEDDEMON نام دارد. این برنامه، یكی از ابزارهای گردآوری اخبار است كه به صورت مستقل كار می‌كند و غیررایگان است.

منبع: سایت تبیان: <http://www.tebyan.net/newindex.aspx?pid=29963>

### مرتبط با رسم گراف در شبیه ساز

**Introduction**

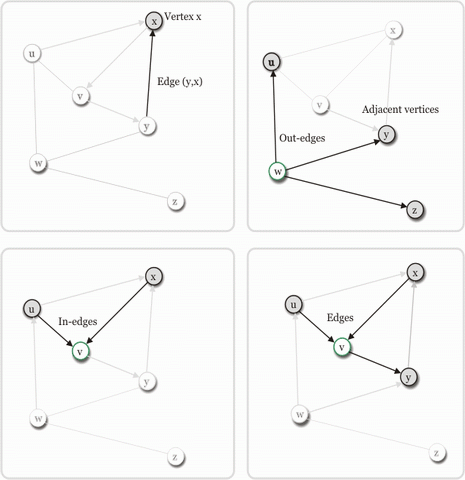
This article presents a **Generic Graph Library**, 100% C#. This library is an attempt to port the Boost Graph Library (BGL) from C++ to C#.

Graph problems arise in a number of situations (more often that you would think): file compilation order, network band-with, shortest path, etc. The library provides the basic data structure to represent vertices, edges and graphs, and also provides **generic** implementation of various graph algorithms such as the depth-first-search, the Dijkstra shortest path, etc.

As the library comes with a full NDoc reference manual, this article will not enter into deep coding details.

**About Graph Theory**

This is a quick remainder about graph theory:



A directed **graph** G=(VxE) consists of a finite set V=V(G) of **vertices** and a finite multi-set E contained in VxV = {(u,v)| u,v in V} of **edges** that are ordered pair of vertices.

If e=(u,v), then e is an **outgoing edge** of u and an **incoming edge** of v. indegree(v) denotes the number ofincoming edges of v and outdegree(u), the number of outgoing edges of u.

Classic graph examples are:

* Transportation network:
  + vertex = crossing,
  + edge = road,
* Internet:
  + vertex = computer,
  + edge = telephone line,
* Source code files:
  + vertex = file,
  + edge = dependency

**What is the BoostGraphLibrary?**

This section discusses aspects about the porting from C++ to C# of the BoostGraphLibrary. The BoostGraphLibrary is a C++ generic graph library that introduces itself as such:

"Graphs are mathematical abstractions that are useful for solving many types of problems in computer science. Consequently, these abstractions must also be represented in computer programs. A standardized generic interface for traversing graphs is of utmost importance to encourage reuse of graph algorithms and data structures. **Part of the Boost Graph Library is an interface for how the structure of a graph can be accessed using a generic interface that hides the details of the graph data structure implementation.** This is an "open" interface in the sense that any graph library that implements this interface will be interoperable with the BGL generic algorithms and with other algorithms that also use this interface. BGL provides some general purpose graph classes that conform to this interface, but they are not meant to be the "only" graph classes; there certainly will be other graph classes better for certain situations. We believe that the main contribution of the BGL is the formulation of this interface."

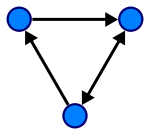
Introduction taken from the [BGL introduction page](http://www.boost.org/libs/graph/doc/).

QuickGraphLibrary has been highly influenced by the BGL implementation and, indeed, a large part of the concept and structure of the library is directly taken from the BGL. Here below, some major points about porting the BGL to C# are discussed.

منبع : http://www.codeproject.com/Articles/5603/QuickGraph-A-100-C-graph-library-with-Graphviz-Sup#xx5110839xx

مرتبط با نوع گراف

گراف جهت‌دار

[](https://commons.wikimedia.org/wiki/File:Directed.svg?uselang=fa)

یک گراف جهتدار

در ریاضیات و به طور خاص در [نظریهٔ گراف](https://fa.wikipedia.org/wiki/%D9%86%D8%B8%D8%B1%DB%8C%D9%87%D9%94_%DA%AF%D8%B1%D8%A7%D9%81)، گراف جهت‌دار یا گراف سودار[[۱]](https://fa.wikipedia.org/wiki/%DA%AF%D8%B1%D8%A7%D9%81_%D8%AC%D9%87%D8%AA%E2%80%8C%D8%AF%D8%A7%D8%B1#cite_note-1) گرافی (مجموعه‌ای از گره‌ها که با یال‌ها به هم متصل شده‌اند) است که در آن به هر یال جهتی نسبت داده شده‌است. به زبان ریاضی، یک گراف جهت‌دار زوج مرتبی به صورت {\displaystyle G=(V,A)} است (گاهی به صورت {\displaystyle G=(V,E)} نیز نمایش داده می‌شود) که در آن

V مجموعه‌ایست که اعضایش را رأس یا گره می‌نامند

A مجموعه‌ای از زوج‌های مرتبی از رأس‌ها است که کمان، یال جهت‌دار، فلش یا گاهی یال نامیده می‌شوند (که در حالت اخیر مجموعهٔ متناظر را به جای A، با E نمایش می‌دهند).

گراف جهت‌دار با گراف معمولی (بدون جهت) متفاوت است: در گراف معمولی، E به صورت مجموعه‌ای از زوج‌های نامرتب از رأس‌ها تعریف شده و معمولاً یال نامیده می‌شود.

یک گراف جهت‌دار ساده نامیده می‌شود اگر هیچ طوقه و یال چندگانه‌ای نداشته باشد (یال‌های چندگانه یعنی یال‌هایی که ابتدا و انتهای یکسانی دارند). در یک گراف چندگانهٔ جهت‌دار یال‌ها تشکیل یک مجموعهٔ چندگانه (به جای مجموعه) از زوج‌های مرتب رأس‌ها می‌دهند و این گراف‌ها می‌توانند طوقه و یال چندگانه داشته باشند (طوقه یالی است که ابتدا و انتهایش رأس یکسانی است). در برخی متون، گراف جهت‌دار (بدون ذکر ویژگی ساده بودن) می‌تواند طوقه، یال چندگانه یا هر دو را داشته باشد.

اصطلاحات پایه‌ای

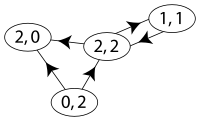
یال {\displaystyle e=(x,y)} دارای جهت از {\displaystyle x} به {\displaystyle y} در نظر گرفته می‌شود. {\displaystyle y} ابتدا و {\displaystyle x} انتهای یال نامیده می‌شود. {\displaystyle y} را رأس مابعد مستقیم {\displaystyle x} و {\displaystyle x} را رأس ماقبل مستقیم {\displaystyle y} می‌نامند. اگر مسیری متشکل از یک یا چند یال رو به جلو از {\displaystyle x} به {\displaystyle y} وجود داشته باشد، به {\displaystyle y} رأس مابعد {\displaystyle x} و به {\displaystyle x} رأس ماقبل {\displaystyle y} می‌گویند. به یال {\displaystyle (y,x)} معکوس یال {\displaystyle (x,y)} می‌گویند.

جهت‌دار کردن یک گراف سادهٔ بدون جهت یعنی نسبت دادن جهت به هر یال آن گراف. هر گراف جهت‌داری که به این طریق ساخته شود گراف جهت‌دار شده (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): oriented graph) نامیده می‌شود. یک گراف جهت‌دار، یک گراف سادهٔ جهت‌دار شده است اگر و تنها اگر نه طوقه داشته باشد و نه حلقهٔ دوتایی (یال چندگانه). [[۳]](https://fa.wikipedia.org/wiki/%DA%AF%D8%B1%D8%A7%D9%81_%D8%AC%D9%87%D8%AA%E2%80%8C%D8%AF%D8%A7%D8%B1#cite_note-3)

گراف جهت‌دار وزن‌دار (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): weighted digraph) گراف جهت‌داری است که به هر یالش وزنی نسبت داده شده‌است، درست مانند گراف بدون جهت وزن‌دار. در نظریهٔ گراف به گراف جهت‌دار با یال‌های وزن‌دار، شبکه (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): network) می‌گویند.

ماتریس مجاورت (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): adjacency matrix) یک گراف جهت‌دار (با طوقه و یال‌های چندگانه) ماتریسی دارای مقادیر صحیح است که سطرها و ستون‌هایش متناظر با گره‌ها هستند و در آن، مؤلفهٔ غیر قطری {\displaystyle a\_{i,j}} تعداد یال‌ها از رأس i به رأس j و مؤلفهٔ قطری {\displaystyle a\_{i,i}} تعداد طوقه‌های رأس i است. ماتریس مجاورت یک گراف جهت‌دار (با یکسان در نظر گرفتن جایگشت‌های سطرها و ستون‌ها) یکتاست. نمایش ماتریسی دیگری برای یک گراف، ماتریس وقوع(به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): incidence matrix) آن است.

درجهٔ ورودی و درجهٔ خروجی

[](https://commons.wikimedia.org/wiki/File:DirectedDegrees.svg?uselang=fa)

یک گراف جهت‌دار که رأس هایش به صورت (درجهٔ خروجی,درجهٔ ورودی) برچسب گذاری شده‌اند.

برای هر گره، تعداد رأس‌های ابتدایی مجاور، درجهٔ ورودی (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): indegree) و تعداد رأس‌های انتهایی مجاور، درجهٔ خروجی (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): outdegree) (در درخت‌ها ضریب انشعاب(به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): branching factor)) نامیده می‌شود.

درجهٔ ورودی با {\displaystyle \deg ^{-}(v)} و درجهٔ خروجی با {\displaystyle \deg ^{+}(v)} نمایش داده می‌شود. رأس با {\displaystyle \deg ^{-}(v)} صفر منبع (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): source) نامیده می‌شود (چون مبدأ همه‌ی یال‌های متصل به خودش است) و به طور مشابه رأس با {\displaystyle \deg ^{+}(v)} صفر چاه (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C):sink) نامیده می‌شود.

رابطهٔ مجموع درجات بیان می‌کند که برای یک گراف جهت‌دار{\displaystyle \sum \_{v\in V}\deg ^{+}(v)=\sum \_{v\in V}\deg ^{-}(v)=|A|\,}

اگر برای هر گرهٔ *v* ∈ *V*، {\displaystyle \deg ^{+}(v)=\deg ^{-}(v)} گراف را یک گراف جهت‌دار متعادل (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): balanced digraph) می‌نامیم. [[۴]](https://fa.wikipedia.org/wiki/%DA%AF%D8%B1%D8%A7%D9%81_%D8%AC%D9%87%D8%AA%E2%80%8C%D8%AF%D8%A7%D8%B1#cite_note-4)

همبندی گراف جهت‌دار

می‌گوییم یک گراف جهت‌دار ضعیفاً همبند است (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): weakly connected) (یا همبند است [[۵]](https://fa.wikipedia.org/wiki/%DA%AF%D8%B1%D8%A7%D9%81_%D8%AC%D9%87%D8%AA%E2%80%8C%D8%AF%D8%A7%D8%B1#cite_note-5)) اگر گراف زمینهٔ آن (که با حذف جهت یال‌ها به دست می‌آید) یک گراف همبند باشد. می‌گوییم یک گراف جهت‌دار قویاً همبند است (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): strongly connected) (یا قوی است) اگر برای هر جفت رأس u و v مسیر جهتداری از u به v و مسیر جهتداری از v به u داشته باشد. مؤلفه‌های قوی (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): strong components)، زیرگراف‌های قویاً همبند ماکزیمال هستند.

انواع گراف‌های جهت‌دار

گراف جهت‌دار G متقارن (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): symmetric) نامیده می‌شود اگر برای هر کمان متعلق به G، معکوس آن کمان نیز متعلق به G باشد. یک گراف جهت‌دار متقارن بدون طوقه معادل یک گراف بدون جهت است که با گذاشتن یک یال به جای هر جفت کمان معکوس هم به دست می‌آید (و بنابراین تعداد یال‌هایش نصف تعداد کمان‌های گراف اولیه است).

یک گراف جهت‌دار سادهٔ بی‌دور

[](https://commons.wikimedia.org/wiki/File:Directed_acyclic_graph_2.svg?uselang=fa)یک گراف جهت‌دار بی‌دور (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): acyclic) یا گراف بی‌دور جهت‌دار، گرافی جهت‌دار است که هیچ دور جهت‌داری ندارد. درخت‌های چندگانه (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): multitrees) (گراف‌هایی که در آن‌ها هیچ دو مسیر جهت‌داری که از یک گره شروع شده‌اند دوباره در یک گرهٔ مشترک تمام نمی‌شوند)، درخت‌های جهت‌دار (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): oriented trees) یا چنددرخت‌ها (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): polytrees) (گراف‌های جهت‌داری که از جهت‌دار کردن یال‌های یک گراف بدون جهت بی‌دور ساخته می‌شوند) و درخت‌های ریشه‌دار (به [انگلیسی](https://fa.wikipedia.org/wiki/%D8%B2%D8%A8%D8%A7%D9%86_%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C): rooted trees) (درخت‌های جهت‌داری که تمام یال‌های درخت بدون جهت زمینهٔ آن به سمت دور شدن از ریشه جهت‌دار شده‌اند) حالت‌هایی خاص از گراف‌های جهت‌دار بی‌دور هستند.

منبع: https://fa.wikipedia.org/wiki/%DA%AF%D8%B1%D8%A7%D9%81\_%D8%AC%D9%87%D8%AA%E2%80%8C%D8%AF%D8%A7%D8%B1

1. پروژه های Big data: پروژه هایی که روی آنالیز انبوهی از اطلاعات پایه گذاری می شوند. [↑](#footnote-ref-1)
2. http://www.infoworld.com/article/2613587/big-data/the-real-story-of-how-big-data-analytics-helped-obama-win.html [↑](#footnote-ref-2)
3. علی جنتی [↑](#footnote-ref-3)
4. Alexa.com [↑](#footnote-ref-4)
5. قابلیتی که می توان با آن بر روی قسمتی از برنامه تغییراتی اعمال کرد و کد نویسی انجام داد. [↑](#footnote-ref-5)
6. خزندگی ربات [↑](#footnote-ref-6)
7. شبیه ساز [↑](#footnote-ref-7)
8. انبوهی از داده ها [↑](#footnote-ref-8)
9. Weight word: الگوریتم پیشرفته برای وزت دادن به کلمات و موضوع بندی آن ها

   مثال: علی مرد خوبی است. 🡸 علی= 10 اجتماعی - مرد = 10اجتماعی، 15 سیاسی، 10 مذهبی – خوب= 5 اجتماعی 🡸 پیام اجتماعی [↑](#footnote-ref-9)
10. این دو سایت متعلق دو [↑](#footnote-ref-10)
11. با این thread که در لغت به معنی گره است می توان یک کد را از حالت یک مسیر اجرا خارج کرد. [↑](#footnote-ref-11)
12. این پروژه در قسمت تلگرام از دیتابیس SQL استفاده می کند و با استفاده از Entity Framework به آن متصل می شود.

    و در قسمت وایبر از دیتابیسی با فایل های تکست استفاده می کند. [↑](#footnote-ref-12)
13. این الگوریتم با بررسی کردن و وجود و عدم وجود کلمات و حروف هایی در پیام و متن و جایگاه آن ها می تواند به میزان تشابه دو متن پی ببرد. [↑](#footnote-ref-13)
14. تشابه یابی شده با استفاده از الگوریتم تشابه یابی پروژه(Findsimilarity) [↑](#footnote-ref-14)
15. امکانی که سازنده های برنامه ها به کاربران می دهند تا درمحدوده ی کارآمادی خاصی توانایی کد نویسی داشته باشند. [↑](#footnote-ref-15)