Tumult

Web Mining 2017

Kewin Dousse - Vincent Robatel

Presentation

1. Context
2. Data
3. Planning
4. Use cases
5. Tools & Techniques

   Demo

6. Conclusion

1 Context

What even IS Discord ?

DISCORD

- VoIP and text chat application
- Free
- Available on most OS
- 45m users in May 2017

Organisation :

- Server
  - Text channel
  - Text channel
  - Voice channel
  - Voice channel

Tumult

Goals :

- Visualize conversations on a server and channel
  - Table of messages
  - Time plot
  - Graph of answers

2

Data

What we worked with

The Discord API

Structure :

- Server
  - Channel
    - Message

Message :


```
{
    "id": "162701077035089920",
    "channel_id": "131391742183342080",
    "author": {},
    "content": "Hey guys!",
    "timestamp": "2016-03-24T23:15:59.605000+00:00",
    "reactions": []
}
```


3

Planning

How did the project happen

Tasks

- Technology choice
- Web app (roads, logic, ...)
- Oauth2 for Discord
- Database script
- Algorithms (word count,...)
- Vis.js and HighStock

4

Use cases

What can I do with it

Features

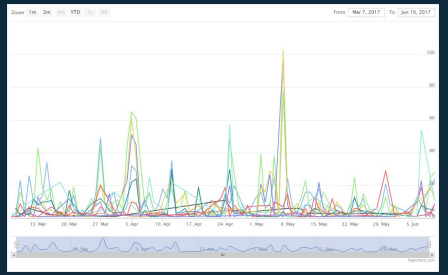
Messages table

Who messages who

Sent / Answered	Nadeko	Blagouz	kurogurochan
Nadeko	1	0	1
Blagouz	0	0	0
kurogurochan	0	0	140

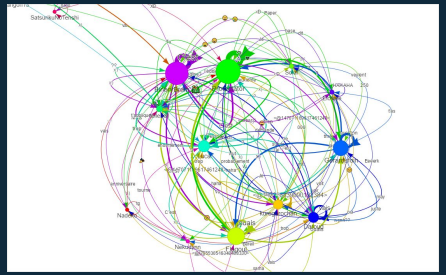
Time plot

Tendencies over multiple days



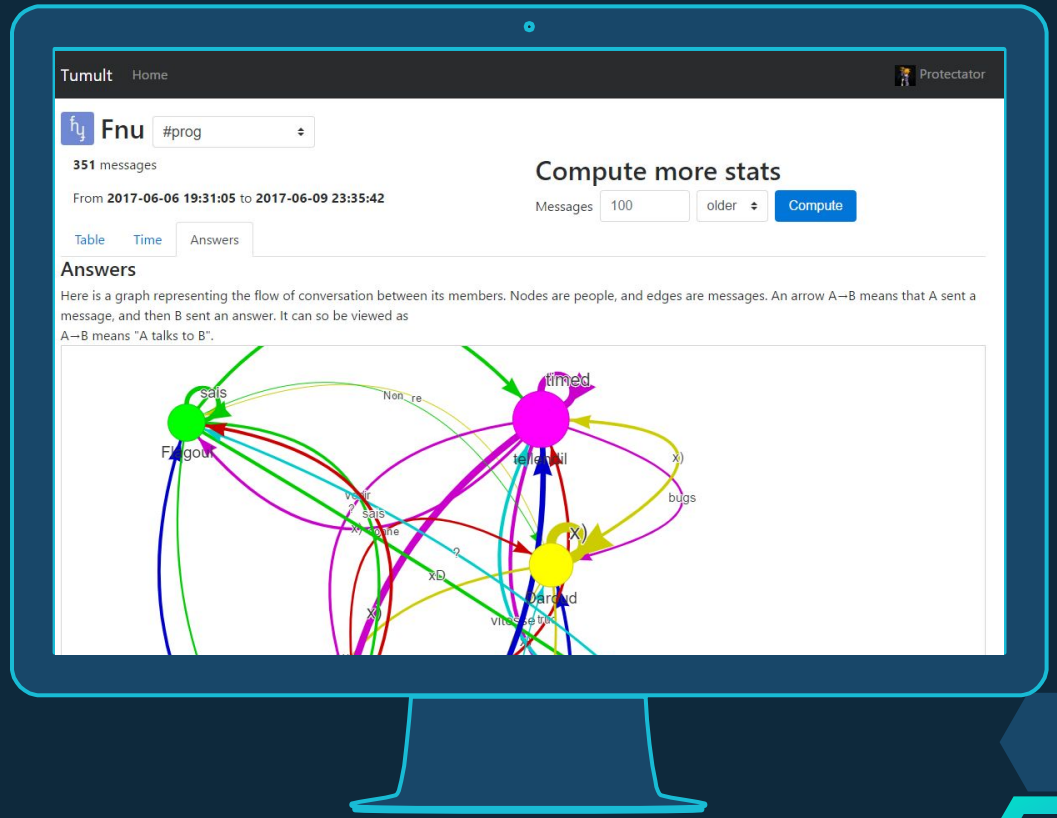
Answers graph

Visualization of the conversation flow



Web app

See the result in your browser !



5

Tools & Techniques

How did we do that

Querying the API

OAuth

Allows the user to connect with his account.

Allows us to retrieve account info, avatar etc.

Limitation : Can’t read messages

User token

User provides this token.

Allows us to get messages from the channels without being registered as a *bot*.

Limitation : Careful with the amount of requests, don’t get banned !

Algorithms

Model of answers system

X : Hi, how’re you ?
Y : Fine !

X → **Y**

X talks to **Y**

Y answers to **X**

Aggregations

To create a table of answers and a time graph :

- Sort & Group by :
  - Time
  - Author
  - Last author
  - Word

Technology Stack

Backend

- Python 3
- Flask + Jinja2
- MySQL

Frontend

- Bootstrap 4
- Vis.js
- HighStock

Demo !

6 Conclusion

Tumult

- Interesting & new data source
- Algorithms to model and aggregate conversations
- Fun results
- Visually interesting

Thanks!

Any questions?

You can find the project at

- github.com/Protectator/Tumult

Credits

Special thanks to all the people who made and released these awesome resources for free:

- Presentation template by SlidesCarnival