

Análise de Redes Regionais da Twitch: Ligações e Padrões de Interação

Trabalho de Grupo realizado no âmbito da Unidade Curricular
de Análise de Redes Avançada do 1º ano do Mestrado em Ciência
de Dados

Diogo Freitas, 104841, MCD-LCD

daafs@iscte-iul.pt

João Francisco Botas, 104782, MCD-LCD

Joao_Botas@iscte-iul.pt

Rebeca Sampaio, 126628, MCD-LCD

rhms01@iscte-iul.pt

Índice

Introdução	1 / 30
O que é a Twitch?	1 / 30
Como funciona?	1 / 30
Impacto da Twitch na era digital	1 / 30
Revisão de Literatura	1 / 30
Entendimento das redes a estudar	3 / 30
Métodos	4 / 30
Tratamento de dados - > https://dev.twitch.tv/docs/api/reference/	4 / 30
Análise Exploratória de dados	5 / 30
Modelling	6 / 30
Evaluation	7 / 30
Deployment	8 / 30
Anexos	8 / 30
Anexo A - Interface Twitch	8 / 30

Introdução

No âmbito da unidade curricular de Análise de Redes Avançada

O que é a Twitch?

Fundada em 2011 e adquirida pela Amazon em 2014, a Twitch é uma plataforma de transmissão ao vivo (*live-streaming*) focada principalmente em videogames, mas abrangendo também outros tipos de conteúdo, como música, desporto, arte e conversas ao vivo. Foi projetada para ser uma plataforma de conteúdo que incluísse torneios de e-Sports, *streams* pessoais de jogadores individuais e programas de entrevistas relacionados a jogos. A página inicial da Twitch atualmente exibe jogos com base na visualização. O espectador típico é do sexo masculino e tem entre 18 e 34 anos.

Como funciona?

A plataforma conta com 4 características que destacam a singularidade e sucesso no universo de transmissões *on-line*:

- Transmissões ao vivo: Os criadores de conteúdo (chamados de *streamers*) transmitem ao vivo enquanto jogam, conversam ou realizam outras atividades;
- Interação em tempo real: A Twitch é conhecida pelo seu *chat* ao vivo, que permite uma interação instantânea entre os *streamers* e a audiência. Adicionalmente, recursos inovadores como mensagens lidas por Inteligência Artificial (IA) aumentam o dinamismo das transmissões, onde mensagens enviadas por utilizadores são convertidas automaticamente em áudio;
- Monetização: Os *streamers* podem ser monetizados por meio de:
 - Inscrições pagas (Subscriptions);
 - Doações de espectadores (Bits e doações diretas);
 - Publicidade (Advertisements de entidades externas exibidos durante as transmissões);
 - Parcerias e patrocínios que os próprios *streamers* têm.
- Programas personalizados de apoio a *streamers*:
 - Afiliado:
 - Parceiro:

Impacto da Twitch na era digital

A Twitch transformou a forma como as pessoas consomem conteúdo ao vivo, permitindo conexões diretas entre criadores e espectadores. Influenciou diretamente o sucesso de vários jogos, como Among Us, Fall Guys, entre outros. Para além disso a plataforma dá voz a criadores de diferentes culturas, promovendo a diversidade. Contudo enfrenta alguns problemas a nível da moderação (apesar de cada criador poder atribuir moderadores à sua escolha para as suas *streams*) tais como discursos de ódio, assédio e conteúdo inadequado.

Revisão de Literatura

Nos últimos anos, a análise de redes sociais tem registado um avanço significativo, graças ao desenvolvimento de técnicas baseadas em matrizes e grafos, apoiadas por ferramentas informáticas. Este progresso tem sido complementado pela aplicação de estatística e matemática, que ajudam a tornar as análises mais objetivas e rigorosas.

O enquadramento teórico das redes sociais privilegia as relações entre indivíduos para compreender a estrutura social, contrastando com as abordagens tradicionais das ciências sociais. Nessas abordagens clássicas, parte-se da definição de categorias preexistentes (como classes sociais, grupos ou organizações), para depois se identificar unidades independentes que são posteriormente agrupadas, com o objetivo de analisar a consistência dos seus comportamentos. Contudo, este método tende a desconsiderar informações importantes resultantes das interações entre as entidades sociais.

Embora grande parte das teorias sociológicas se foque nas relações entre atores, a análise de redes sociais destaca-se por introduzir ferramentas técnicas que permitem verificar empiricamente hipóteses teóricas sobre a natureza das relações e a estrutura das redes.

O significado atribuído à análise de redes sociais é alvo de alguma ambiguidade. Estas incertezas decorrem da diversidade de interpretações e abordagens existentes em diferentes disciplinas e correntes, que conferem ao conceito de rede múltiplos significados, muitas vezes contraditórios, dificultando a sua clarificação.

Apesar dos avanços na área, a análise de redes sociais ainda está amplamente associada a um grupo restrito de cientistas sociais que utilizam uma linguagem técnica muito específica. Esta característica pode representar um entrave para outros investigadores, particularmente para aqueles que estão mais habituados a trabalhar com abordagens baseadas na lógica dos atributos para estudar fenómenos sociais.

Nesta linguagem técnica, matrizes e grafos destacam-se como ferramentas fundamentais para mapear e ilustrar as interações entre indivíduos, grupos e organizações. No entanto, como apontam Alejandro e Norman (2005), as particularidades da análise de redes sociais tornam inadequadas muitas das ferramentas estatísticas tradicionalmente utilizadas para análises sociais.

(PARTE DO ARTIGO https://dspace.uevora.pt/rdpc/bitstream/10174/12831/1/20881-41419-1-PB_FINAL.pdf)

Entendimento das redes a estudar

Para a realização deste projeto foram extraídas as redes da Twitch do portal de [Stanford](#). Estas redes fazem parte do [MUSAE](#), uma coleção de *datasets* projetada para estudar e auxiliar pesquisas na área de ciência de redes, *Machine Learning* em grafos, análise de redes sociais e detecção de comunidades. Estes grafos são criados a partir de redes sociais como Facebook, GitHub, Twitch, entre outras.

No caso da Twitch, dão acesso a redes de 6 línguas diferentes que contêm as informações *user-user* entre *streamers* da Twitch, recolhidas em Maio de 2018. Ou seja, temos 6 redes semelhantes que são estáticas, porque foram extraídas num determinado momento do tempo; não direcionadas, pois a ligação entre dois *streamers* é mútua. Como primeiro passo para a análise, é necessário compreender as redes iniciais sem que tenham sido efetuadas quaisquer alterações.

Tal como foi mencionado, eram fornecidas redes com diferentes idiomas, sendo estas DE, EN, ES, FR, PT-BR e RU, que correspondem a *streams* onde a língua principal é alemão, inglês, espanhol, francês, português (do Brasil) e russo, respetivamente. Os nodos representam um *streamer* de uma das 6 línguas especificadas, enquanto as ligações representam as amizades entre esses *streamers*.

A informação do número de nodos e de arestas de cada uma destas redes encontra-se na Tabela 1.

Tabela 1: N° de nodos e arestas da rede inicial

Língua da Rede	N° de nodos	N° de arestas
DE	9498	153138
EN	7126	35324
ES	4648	59382
FR	6549	112666
PT-BR	1912	31299
RU	4385	37304

É evidente que, observando para a Tabela 1, as redes têm alguma disparidade entre elas, tendo em conta o número de nodos e arestas. Isto deve-se sobretudo ao facto de

Para cada uma das línguas eram fornecidos dois ficheiros, um com a informação sobre os nodos, com informação subjacente aos nodos/*streamers* (*musae_XXX_target.csv*) e outro com a informação das arestas entre nodos/ligações entre *streamers* (*musae_XXX_edges.csv*), onde “XXX” corresponde ao código da língua apresentado acima.

No ficheiro com a informação dos nodos tínhamos acesso às seguintes variáveis correspondentes ao *streamer*:

- **id**: Id único associado à conta da Twitch do *streamer*. Este valor corresponde mesmo a um id verdadeiro da própria Twitch e será dado mais importância na **COLOCAR LINK SECÇÃO**
- **days**: há quantos a conta foi criada;
- **mature**: variável que indica se o *streamer* exibe conteúdo adulto nas suas *streams*;
- **views**: quantas views o nodo/*streamer* teve na totalidade do seu canal;
- **partner**: se o *streamer* é um utilizador parceiro da Twitch;
- **new_id**: id diferente e único que serve para fazer correspondência de arestas entre nodos.

Métodos

Tratamento de dados - > <https://dev.twitch.tv/docs/api/reference/>

```
1  {
2    id: 141981764,
3    login: "davyjones",
4    display_name: "DavyJones",
5    type: "",
6    broadcaster_type: "partner",
7    description: "Supporting third-party developers",
8    profile_image_url: "https://encurtador.com.br/jS9rY",
9    offline_image_url: "https://encurtador.com.br/ZGdUa",
10   view_count: This field has been deprecated,
11   email: "not-real@email.com",
12   created_at: 2016-12-14T20:32:28Z
13 }
```

JSON

Listagem 1: GET <https://api.twitch.tv/helix/users>

```
1  {
2    broadcaster_id: 141981764,
3    broadcaster_login: twitchdev,
4    broadcaster_name: "TwitchDev",
5    broadcaster_language: "en",
6    game_id: 509670,
7    game_name: "Science & Technology",
8    title: "TwitchDev Monthly Update // May 6, 2021",
9    delay: 0,
10   tags: ["DevsInTheKnow"],
11   content_classification_labels: ["Gambling", "DrugsIntoxication", "MatureGame"],
12   is_branded_content: False
13 }
```

JSON

Listagem 2: GET <https://api.twitch.tv/helix/channels>

EXPLICAR NOVAS VARS

Com os dados extraídos foi necessário verificar algumas incongruências que pudessem existir relativamente a valores omissos. Vê-se que alguns nodos não têm username, ...

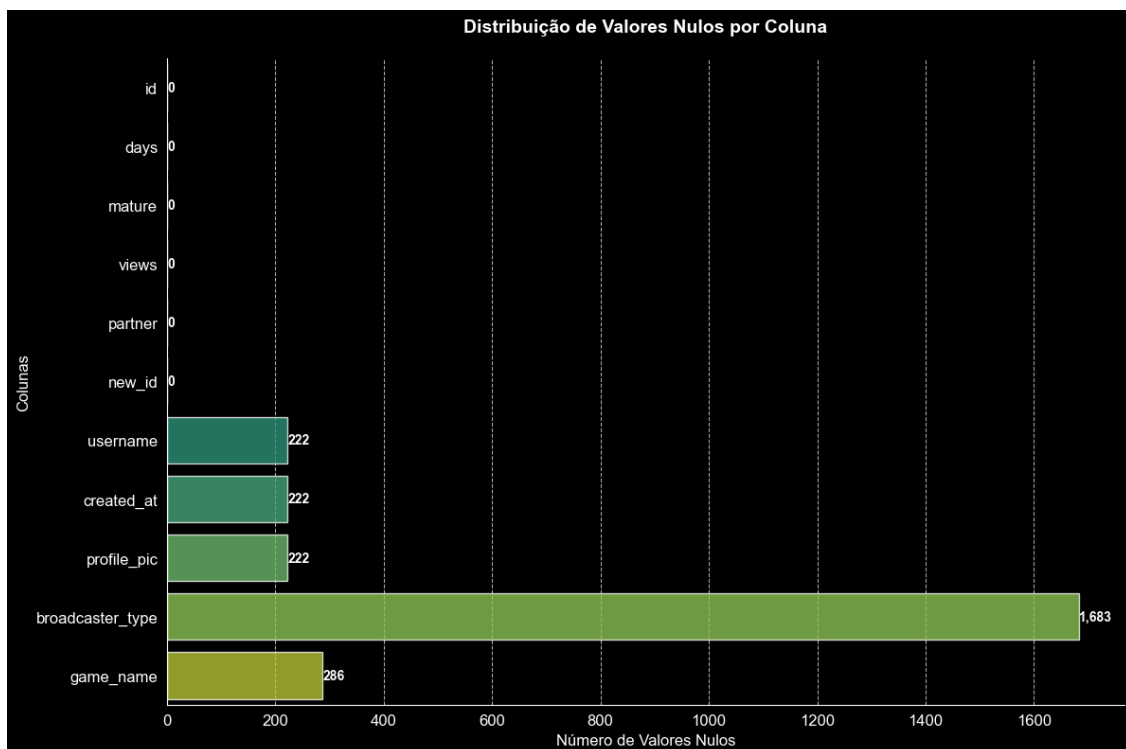


Figura 1: Distribuição de valores omissos para a rede RU - Rússia

Análise Exploratória de dados

Após serem tratados todos os dados, segue-se a análise dos dados recolhidos. Nesta secção iremos dar um maior destaque à rede de ??? porque blablabla. Mais à frente iremos comparar redes de outros idiomas e tirar conclusões mais robustas e sustentadas.

Modelling

Evaluation

Deployment

Anexos

Anexo A - Interface Twitch