Pontifícia Universidade Católica de Campinas Ciência de dados e Inteligência Artificial

Grafos e Redes Sociais

Trabalho Final

THE WITCHER

UMA ANÁLISE DA REDE SOCIAL AOS OLHOS DO R

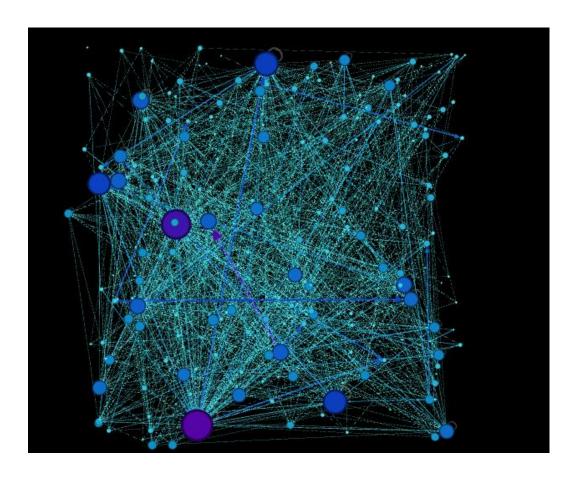
Introdução4	
Coleta de Dados e Preparação da Rede	
Análise e Discussão dos Dados	
Um pouco mais sobre os protagonistas	2
Alteração da Rede ao Longo dos Livros	,
Conclusão	2
Bibliografia	3

Introdução

The Witcher, universo criado pelo autor Andrzej Sapkowski, se passa em um mundo de fantasia medieval e segue a história de Geralt de Rívia, um assassino impiedoso e de sangue-frio treinado, desde a infância, para caçar e eliminar monstros, além de ser um dos últimos bruxos restantes na Terra. Ele busca encontrar seu lugar em um mundo onde as pessoas muitas vezes são mais perversas que as criaturas selvagens, um mundo em que nem todos os que parecem monstros são maus nem todos os que parecem anjos são bons.

Neste trabalho foi realizado uma análise da rede social do universo de The Witcher utilizando o RStudio (um software livre de ambiente de desenvolvimento integrado para R) em conjunto com a biblioteca "Igraph", utilizada para a criação e manipulação de grafos. Também foi utilizado o software Gephi para que se possa ter uma melhor visualização dos grafos construídos.

O principal objetivo foi analisar as relações entre os personagens na rede como um todo, percebendo com quem e com qual intensidade eles interagem, além de expor o desenvolvimento dos mesmos durante o andar da série de sete livros.



Coleta de Dados e Preparação da Rede:

Para realizar a construção dos grafos utilizados neste trabalho foi necessário buscar um dataset já existente no site Kaggle, uma comunidade online de cientistas de dados e profissionais de aprendizado de máquina utilizada para explorar e publicar conjuntos de dados.

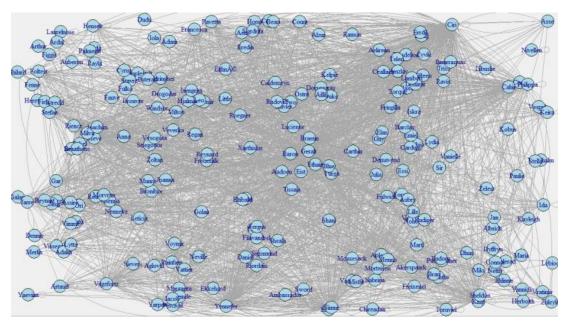
Nesse site foi possível encontrar um arquivo .csv criado por Ada Adasivan que continha as conexões entre os personagens, composta por duas colunas (Sourse e Target) e classificadas por livro do 1 ao 7, sendo que cada interação possui um peso determinado pelo número de vezes que ela ocorre em seu respectivo livro.

Tendo em vista que obtivemos apenas a lista de arestas do nosso grafo, tivemos que unir através do Excel as colunas contendo as conexões e eliminar os dados duplicados, resultando assim em uma única coluna composta pelos personagens.

Para realizar a análise dos dados coletados, utilizamos diversos conceitos advindos da teoria dos grafos como: grau, intermediação de vértice, distância, densidade, centralidade de autovetor, subgrafos, cliques, vizinhança e comunidades.

Análise de Dados e Discussão:

Primeiramente foi criado um grafo utilizando o arquivo das arestas já obtidas no site Kaggle e o arquivo criado dos vértices.



Dentro do universo dos livros de The Witcher, foi identificado 224 personagens (que trataremos neste trabalho como nós ou vértices) e 2600 conexões (arestas) entre eles, sendo o primeiro livro o que apresenta maior número de personagens (188) e o sétimo o maior número de interações entre os personagens (622). Tendo em vista que um mesmo nó pode aparecer em mais de um livro, precisaremos interpretar com cuidado alguns dos resultados obtidos.

- Graus:

O grau é uma métrica interessante de ser usada, visto que ela indica o número de arestas a qual um determinado vértice está conectado.

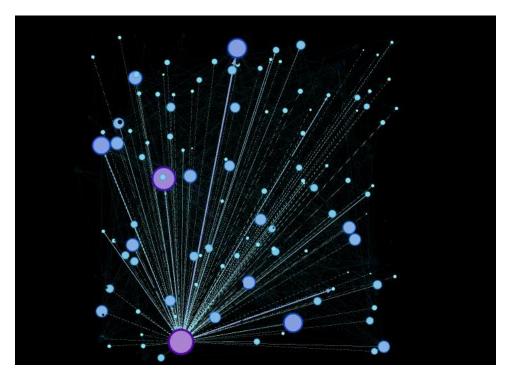
Na rede social em análise, a média dos graus é 23.2, ou seja, cada personagem tem uma média de 23 interações ao longo dos livros.

```
sort(grau, decreasing = TRUE)[1:10]
Geralt Ciri Yennefer Dandelion Emhyr Philippa Triss King Vilgefortz Calanthe
404 331 208 188 150 125 108 103 102 98
```

Na imagem acima é possível ver os dez nós com maior grau, tendo Geralt como o vértice de maior grau com 404 arestas conectadas a ele. Este número pode ser interpretado como o número de personagens que Geralt interagiu na

totalidade dos livros, e ele é maior que o número total de nós pelo fato de que existem personagens que interagiram com ele em mais de um livro. Nas análises subsequentes, serão vistas diversas outras métricas em torno dos quatro personagens principais deste universo, sendo eles: Geralt, Ciri, Dandelion e Yennefer.

Para uma melhor visualização, foi usado o Gephi para mostrar de uma forma mais clara todas as conexões de Geralt.



- Intermediação de Vértice:

A intermediação de vértice pode ser entendida como o quanto um nó é utilizado como ponte ao longo do caminho mais curto entre dois outros vértices, sendo uma medida para quantificar o controle de um ser humano sobre a comunicação entre outros dois seres humanos numa rede social.

Novamente nossos quatro personagens aparecem com os números mais altos, indicando que eles são indispensáveis para a trama dos livros, atuando como ligações entre diversos personagens.

```
> sort(intermediacao_vertice, decreasing = TRUE)[1:10]
Geralt Ciri Yennefer Dandelion Emhyr Jarre King Philippa Count Falka
9161.9320 4729.3658 1635.8536 1305.1074 1063.1996 980.3440 816.3794 526.4830 465.0366 447.5597
```

- Distância:

O personagem Voymir é um cavaleiro do rei Ervyll, que ajudou Ciri algumas vezes a escapar de algumas enrascadas. Porém caso ele se voltasse contra Ciri e resolvesse sequestrá-la, precisaria realizar o seguinte caminho:

```
> shortest_paths(grafo, "Voymir", "Ciri")
$vpath
$vpath[[1]]
+ 5/224 vertices, named, from f26de90:
[1] Voymir Sigismund Lucienne Rayla Ciri
```

ganhando primeiramente a confiança de Sigismund, em seguida de Lucienne e Rayla, para que finalmente conseguisse alcançá-la. Pôde-se chegar a essa conclusão através da análise da distância entre esses personagens.

A distância entre dois vértices refere-se à quantidade de arestas que é preciso percorrer para que se chegue do nó de origem ao nó de destino. A distância média encontrada nesse grafo foi de 2.5, mostrando que nem todos os personagens dos livros são apresentam conexões diretas entre si, necessitando, assim, de uma intermediação para que se conectem à determinadas pessoas.

```
> mean_distance(grafo)
[1] 2.504107
```

Existe uma teoria, proposta pelo psicólogo Stanley Milgram, que diz que são necessários no máximo seis laços de amizade para que duas pessoas quaisquer estejam ligadas. Esta teoria pôde novamente ser testada pela análise do diâmetro de nossa rede, necessitando seguir o seguinte caminho para que a personagem Joanna se encontrar com Voymir (sendo esses dois personagens os mais distantes entre si de todo o grafo):

```
> get_diameter(grafo)
+ 8/224 vertices, named, from f26de90:
[1] Joanna Carthia Foltest Geralt Toruviel Lucienne Sigismund Voymir
```

- Densidade:

A densidade de um grafo é a razão entre a quantidade de arestas do mesmo e a quantidade de arestas caso ele fosse um grafo completo (grau máximo em todos os vértices, que ocorreria se todos os personagens se conhecessem) com a mesma quantidade de vértices.

Neste grafo, a densidade obtida foi de aproximadamente 0.1, mostrando novamente que nem todos os personagens possuem afinidade entre si.

```
> edge_density(grafo)
[1] 0.1040999
```

- Centralidade de Autovetor:

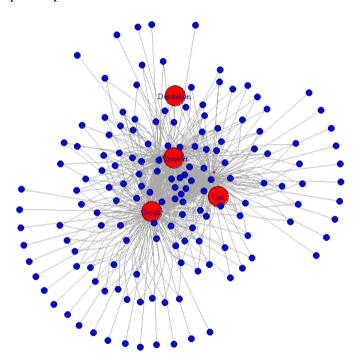
A centralidade de autovetor é uma métrica que indica a influência de um nó numa rede. Ela atribui pontuações relativas a todos os nós da rede, baseada no conceito de que as ligações para os nós de alta pontuação contribuem mais para a pontuação do nó em questão do que ligações de baixa pontuação. Em termos mais simples, quanto mais pessoas de influência alguém está ligado, maior será sua influência na rede. Através de uma manipulação matemática é possível transformar os valores obtidos em porcentagem.

Os números obtidos reforçam o fato de Geralt, Ciri, Dandelion e Yennefer serem os personagens pela qual a trama gira, sendo de longe os mais influentes dessa rede.

- Vizinhança e Subgrafos:

O termo vizinhança, como o próprio nome sugere, refere-se aos vértices que estão conectados em torno de um determinado vértice, podendo especificar a ordem de vizinhança que se deseja obter. Com isso, conseguimos descobrir que Geralt, Ciri, Dandelion e Yennefer estão conectados respectivamente com 116, 89, 60 e 70 personagens únicos.

A imagem a seguir mostra um subgrafo com todas as conexões de nossos personagens principais.



- Cliques:

Os cliques podem ser descritos de jeito informal como os subgrafos em que cada vértice possui conexão com todos os outros, ou seja, todos os vértices desse subgrafo possuem grau máximo. Como podemos verificar na imagem abaixo, o maior número de nós que um clique pode ter neste grafo é 10, havendo 3 cliques de maior largura.

```
> largest.cliques(grafo)
[[1]]
+ 10/224 vertices, named, from clafb97:
[1] Margarita Yennefer Sabrina Emhyr Philippa Ciri Fringilla Dandelion Geralt King

[[2]]
+ 10/224 vertices, named, from clafb97:
[1] Margarita Yennefer Sabrina Emhyr Philippa Ciri Fringilla Dandelion Geralt Triss

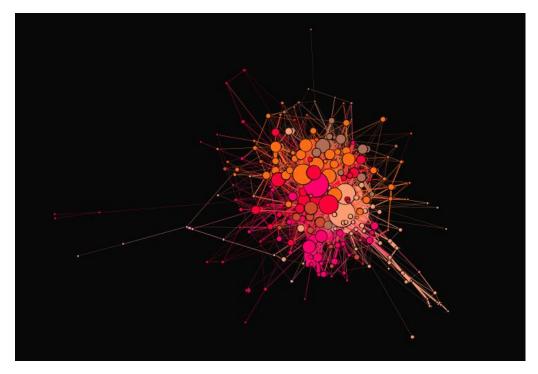
[[3]]
+ 10/224 vertices, named, from clafb97:
[1] Fringilla Geralt Emhyr Philippa Ciri Dandelion Yennefer Sabrina Vilgefortz Triss
```

Caso precisássemos escolher 10 personagens dos livros para compor uma esquipe, poderíamos utilizar um desses 3 cliques obtidos para realizar a seleção. Dessa forma teríamos uma equipe mais unida, visto que todos os dez integrantes se conheceriam, assim tendo mais desempenho em uma dada tarefa.

- Comunidades:

Comunidades, também chamadas de clusters ou agrupamentos, são grupos de vértices que tem grande probabilidade de compartilhar propriedades comuns ou ter papéis semelhantes na rede.

Para a separar nosso grafo em comunidades, utilizamos o cluster_walktrap, que funciona como um algoritmo de passeio aleatório, obtendo assim 78 grupos diferentes.



Dentro da comunidade é possível analisar a sua modularidade, que mede a densidade das ligações internas a um grupo (quanto mais próximo de 1 a modularidade estiver, mais homogênea é a rede). A modularidade obtida nesse grafo foi de aproximadamente 0,2.

Apesar de existirem quase 80 grupos diferentes, apenas alguns deles possuem relevância nessa rede, por isso utilizamos alguns códigos para filtrar quais são os principais grupos (8, 2, 1, 5 e 13) e seus membros.

```
> sort(lista, decreasing = TRUE)[1:5]
membros_comunidade1
    8    2    1    5    13
51    35    13    8    5
```

A análise dos membros dos dois principais grupos (8 e 2) nos permitiu concluir que os integrantes do primeiro são principalmente os protagonistas deste universo e o segundo é composto pelos antagonistas.

Um pouco mais sobre os protagonistas:

- Ciri:

A garota, cujo nome completo é Cirilla Fiona Elen Riannon, é sucessora ao trono de Cintra por ser neta de Calanthe, a rainha. Filha de Pavetta e Duny, ela é criada pela avó após os pais terem sido dados como mortos em um estranho acidente de barco. Inicialmente, Ciri é bastante mimada por ter sido criada com tudo do bom e do melhor, mas rapidamente ela precisa aprender como funciona o mundo real quando seu reino entra em guerra contra o poderoso império de Nilfgaard. Sua situação fica ainda mais extrema quando, durante um impiedoso ataque, Calanthe se torna vítima da violência, e Ciri é perseguida por uma figura de armadura negra.

Após a queda de Cintra, os caminhos de Ciri e Geralt se cruzam novamente, e o bruxo percebe a força do destino em sua vida, ainda que de forma relutante. Para protegê-la, ele a leva para a fortaleza de Kaer Morhen, e passa a treiná-la para eventualmente se tornar uma bruxa.

Yennefer:

Yennefer de Vengerberg, nascida na capital de Aedirn durante o Belleteyn de 1173, é uma poderosa feiticeira e o grande amor do bruxo Geralt de Rívia, além de uma figura materna para Ciri e uma amiga próxima de Triss Merigold.

Grã-mestra de Magia, Yennefer era famosa e muito respeitada, sendo a mais jovem membra do Conselho Supremo.

A pedido de Geralt, Yennefer se tornou tutora de Ciri no intuito ensiná-la sobre magia para controlar seus poderes durante sua estada no templo de Melitele, em Ellander. Com a convivência, uma forte ligação entre as duas se fez e Yennefer passou a considerar Ciri como sua protegida e filha. Ciri também partilhava do sentimento e por diversas vezes chamou a feiticeira de "mãe".

- Dandelion (Yasker):

Julian Alfred Pankratz, visconde de Lettenhove, melhor conhecido como Jaskier (Dandilion), é um poeta, menestrel, bardo e amigo próximo de Geralt de Rívia, o seguindo em várias de suas aventuras

Durante suas viagens com Geralt, Jaskier começou a escrever suas memórias, intitulado "Meio século de Poesia".

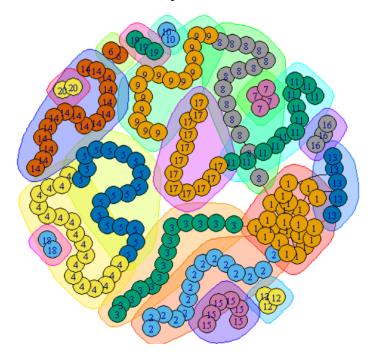
Alteração da Rede ao Longo dos Livros:

Para fazer a análise do desenvolvimento da rede ao longo dos livros, tivemos que filtrar a lista de arestas e criar uma nova lista de vértices para cada livro, construindo grafos para todos eles.

- Livro 1 (O Último Desejo):

O Último Desejo é o primeiro livro da série de livros de Andrzej Sapkowski referente ao universo de The Witcher. Não faz parte da Saga Witcher que começa com o terceiro livro, O Sangue dos Elfos, é uma coletânea de contos passados antes da saga em que acompanhamos Geralt de Rívia perambulando de povoado em povoado oferecendo seus serviços.

No grafo do primeiro livro temos 188 nós e 208 arestas, sendo o livro que possui a maior quantidade de personagens. Por termos poucas conexões em relação ao número de personagens, é possível deduzir que teremos grupos muito fechados e com poucas interações entre personagens de diferentes grupos. Isso fica evidente ao plotarmos a comunidade.



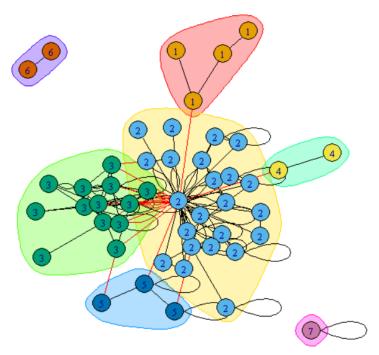
Esta diversidade foi causada pelo fato do primeiro livro ser separado em vários contos relativamente independentes, obtendo assim uma grande variedade de grupos fortes nesse grafo.

Aqui temos Geralt como personagem mais importante e não há sinal dos outros três protagonistas.

- Livro 2 (A Espada do Destino):

Enquanto no livro um, "O Último Desejo", fomos introduzidos a um Geralt mais introspectivo e metódico, no livro dois somos apresentados a um personagem já mais amadurecido e com menos medo de demonstrar os seus sentimentos, com falas mais acaloradas e trejeitos menos robóticos. Seu crescimento gradual, do começo ao fim do livro, é perceptível e feito de maneira primorosa, levando-nos a conhecer melhor o bruxo e sua vida. Assim como o primeiro, este livro é composto por uma coletânea de contos sobre as aventuras de Geralt.

Diferentemente do grafo anterior, neste temos quase o triplo de conexões em relação ao número de vértices (48 nós e 159 arestas), mostrando que, apesar de também ser um livro de contos, as histórias serão mais interdependentes, se conectando umas com as outras.

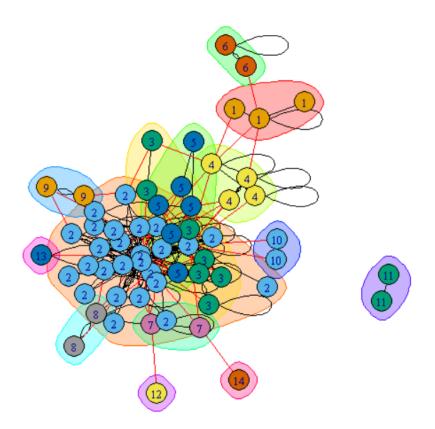


Ao analisar os membros desta comunidade, fica evidente que o grupo 3 refere-se aos membros da realeza de Cintra ou que tem envolvimento com a mesma.

Neste livro Geralt continua sendo o mais influente da rede, porém vemos que Yennefer desempenha um papel mais bem definido, mesmo que ainda seja uma personagem secundária.

- Livro 3 (O Sangue dos Elfos):

O livro três, com 61 personagens e 300 interações entre eles, inicia-se após os eventos que fecharam A Espada do Destino. O bruxo Geralt de Rívia, depois de ter encontrado a criança que lhe foi prometida por um serviço, decide adotá-la e treiná-la nas artes de seu ofício. Para isso, ele leva a menina, Cirilla, até a fortaleza de Khaer Morhan, onde a submete a exercícios intensos. Enquanto isso, assassinos, feiticeiros e reis iniciam uma busca incansável pela criança, visto que ela é uma princesa e a única pretendente viva ao trono do reino de Cintra, destruído pela guerra.

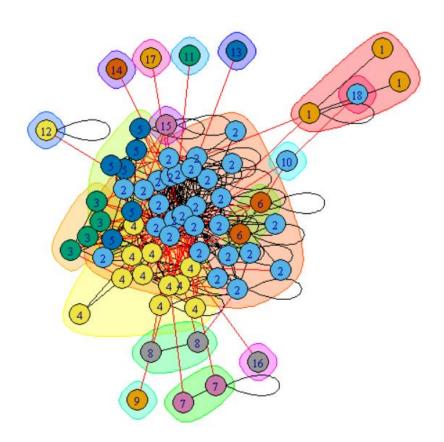


Nesta comunidade temos no grupo 2, com exceção de Rience (um poderoso mago de fogo que busca capturar Ciri), os aliados dos protagonistas, que os ajudam no andar da trama.

Aqui percebemos que Ciri aparece com força na história pela primeira vez e vemos que Dandelion, apesar de já ter acompanhado Geralt em algumas aventuras em momentos anteriores, só demonstra relevância mesmo neste livro.

- Livro 4 (Tempo do Desprezo):

Contendo 68 personagens e 427 arestas, Tempo do Desprezo é o quarto livro da saga do bruxo Geralt de Rívia. Geralt lutou contra monstros e demônios por todo o país, mas até ele pode não estar preparado para o que está acontecendo com seu mundo. Há intrigas, divergências e rebeliões por todo lado. Os Elfos e outros seres não humanos vivem sob repressão há décadas. Os Magos brigam uns com os outros, alguns a soldo dos reis, outros simpatizantes dos elfos. E, nesse cenário de medo e desprezo, Geralt e sua amante Yennefer precisam proteger Ciri, herdeira órfã e procurada por todos os lados. Ela tem o poder de salvar o mundo ou, talvez, acabar com ele.



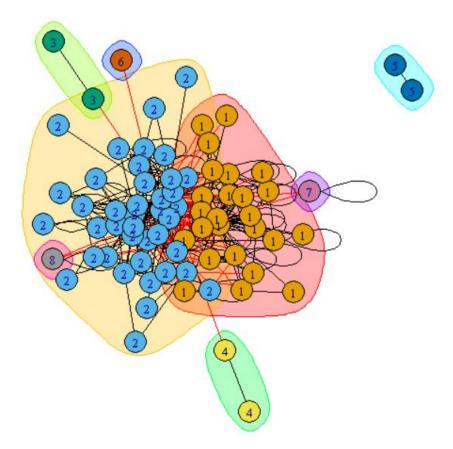
Em meio a esse caos que ocorre no livro 4, praticamente todos os magos e feiticeiros se encontram no grupo 2 dessa comunidade.

No quarto livro da saga é notável o desenvolvimento que Ciri e Yennefer apresentam, sendo o destaque do livro. Isso sucede por conta do caminho que a história toma, aproximando as duas personagens em uma relação quase de mãe e filha.

- Livro 5 (Batismo de Fogo):

No quinto livro a Irmandade dos Magos foi esfacelada por um golpe e Geralt ficou seriamente ferido. Guerras devastam todos os territórios e o futuro da magia está ameaçado. Mas as feiticeiras que sobreviveram estão determinadas a protegê-lo. É um momento conturbado, e ainda é preciso encontrar a garota Ciri, herdeira do trono de Cintra, que desapareceu. Até que surge o rumor de que ela está na corte de Nilfgaard e vai se casar com o imperador. Mesmo ferido, Geralt parte em sua missão de resgatá-la.

O grafo apresenta 74 nós e 450 vértices, contendo apenas dois grupos relevantes nessa comunidade.

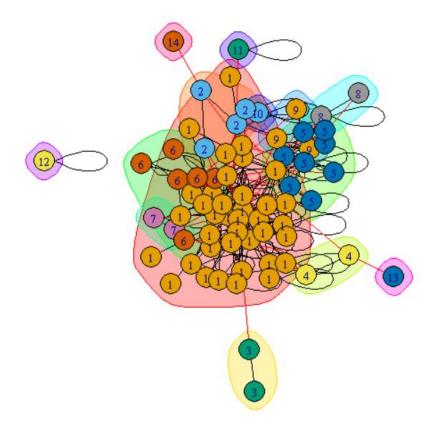


Tendo em vista que Geralt e Ciri foram separados em decorrência dos acontecimentos do livro anterior, eles ficaram em grupos diferentes. Geralt está no grupo 1 e Ciri no 2, junto com membros importantes do Império Nilfigaardiano.

Em decorrência do caos que está ocorrendo, temos uma dispersão nas histórias e aparecem outros personagens como destaque.

- Livro 6 (A Torre da Andorinha):

Com 72 personagens e 337 conexões, o penúltimo livro acompanha o bruxo que segue em direção ao sul, convencido de que Ciri vai se casar em breve com o imperador de Nilfgaard. Os assassinos impiedosos de Nilfgaard já estão seguindo os rastros de Ciri, que, gravemente ferida, se recupera no casebre do eremita Vysogota, perdido no meio de um pantanal.



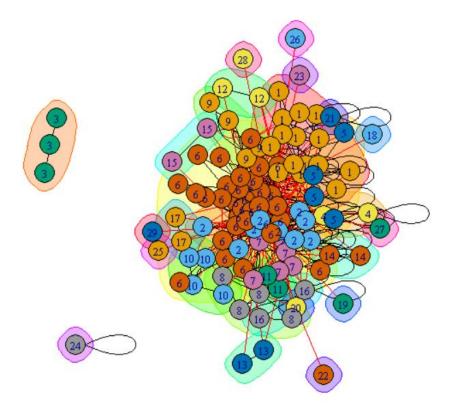
Nessa comunidade percebemos que o grupo 5 é composto majoritariamente por reis e rainhas.

Aqui vemos que Geralt volta a ser o foco do livro, entretanto Ciri apresenta um papel quase tão significativo quanto ele, tendo em vista que ambos possuem uma centralidade de autovetor relativamente semelhante. Em comparação aos outros personagens.

- Livro 7 (A Senhora do Lago):

Neste livro, enquanto Yeneffer permanece prisioneira e Geralt passa o inverno sob o charme de Fringilla, Ciri é projetada em um mundo paralelo ao penetrar na Torre da Andorinha. Agora, sob o domínio do elfo Avallac'h, ela só poderá ser livre se aceitar dar um filho ao rei dos Amieiros.

O último livro da saga de The Witcher contém 110 personagens e 662 arestas, sendo o que possui o maior número de interações entre os personagens.



Nesta comunidade percebeu-se que o grupo 2, com exceção de Devemend, é composto de apenas de feiticeiras.

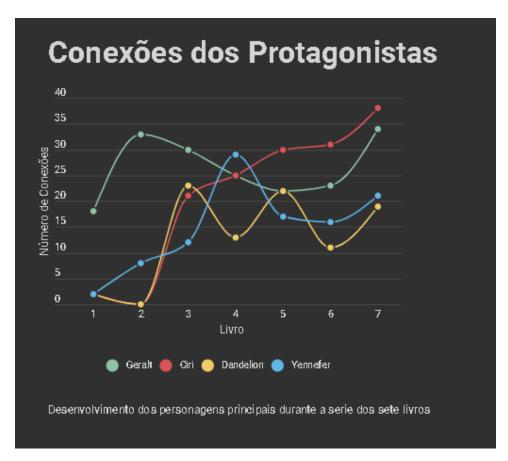
Tendo em vista a diferença nos valores de Ciri e Geralt, podemos afirmar que Ciri é de longe a personagem mais relevante na história do último livro dessa saga.

```
> sort(centralidade_7$vector, decreasing = TRUE)[1:10]
Ciri Geralt Yennefer Emhyr Dandelion Vilgefortz Triss Count Kelpie Fringilla
1.0000000 0.6158068 0.5164797 0.4856380 0.4518467 0.3224515 0.3072011 0.2768882 0.2747244 0.2676987
```

- Concluindo...

Para melhorar a compreensão dos dados obtidos ao analisar cada livro individualmente, construímos um gráfico mostrando esse desenvolvimento percebido em cada personagem ao longo das histórias.

Usamos como métrica de comparação o número de vizinhos que cada um dos personagens principais teve em cada um dos livros, obtendo assim o seguinte gráfico:



Percebemos através dele que nos dois primeiros livros Geralt teve um maior destaque, sendo o momento em que somos introduzidos ao universo do bruxo e suas aventuras iniciais através de contos separados. No segundo livro vemos que Yennefer começa a desenvolver um certo papel na história, porém a personagem ainda não é aprofundada. Podemos observar também que as personagens Ciri e Dandelion só começaram a ganhar relevância depois do terceiro livro, onde eles realmente são apresentados ao público com mais detalhes, já a personagem Yennefer ganha um grande destaque apenas no quarto livro, onde são aprofundadas as relações dela com Ciri. Nos últimos livros acompanhamos a jornada de Ciri para reencontrar seus amigos, onde ela se torna o foco da saga até o fim.

Além das métricas já utilizadas, decidimos analisar também o peso das ligações de Geralt e a média das distâncias dele com os outros personagens ao decorrer das histórias e percebemos que no primeiro livro o personagem tinha na maior parte de suas ligações um baixo peso (o que significa que ele não possuía tanta intimidade com os personagens) e tinha uma distância média de 8.8, já no livro três ele teve mais interações de alto peso e uma distância média de 1.6 aproximadamente, evidenciando que no começo de sua jornada ele era frio e distante dos outros, porém após conhecer melhor nossos outros protagonistas o bruxo acaba amolecendo e passa a desenvolver uma melhor afinidade com outros personagens, se importando mais com eles e amadurecendo nesse sentido.

Conclusão:

Através da análise de diversas métricas de rede social criada podemos perceber que os principais personagens deste universo são "Geralt", "Ciri", "Yennefer" e "Dandelion", verificando também a alteração da influência de cada um ao decorrer dos livros.

Além disso, conseguimos separar essa rede em dois grupos principais: os dos protagonistas e dos antagonistas da séria. Também analisamos as interações ao longo das histórias e conseguimos encontrar padrões que se ligam com os fatos ocorridos nos livros.

As aventuras de The Witcher são vastas, sendo compostas por livros, séries, animações, jogos, entre outros. A proposta deste projeto foi realizar uma análise apenas do conjunto dos livros em uma perspectiva mais generalizada, correlacionando as informações obtidas pelos grafos à apenas alguns aspectos das histórias e assim não se preocupando tanto com os pormenores de cada obra.

Concluindo, com esse trabalho foi possível conhecer mais sobre esse universo mesmo sem precisar ler qualquer um dos livros, mostrando um pouco do potencial dos grafos para a análise de redes sociais, sejam elas reais ou fictícias.

Bibliografia:

https://www.kaggle.com/datasets/avasadasivan/witcher-network

https://legadoplus.com.br/the-witcher-confira-o-resumo-completo-da-historia-da-serie-do-netflix/

https://tionitroblog.wordpress.com/2017/03/22/the-witcher-vols-1-a-7-resenha-de-todos-os-sete-livros-da-saga-do-bruxo-geralt-de-rivia-nitroleituras-fantasia

https://witcher.fandom.com/wiki/Witcher Wiki

https://igraph.org/r/doc/cluster_walktrap.html