

UNIVERSIDADE PRESBITERIANA MACKENZIE
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E COMPUTAÇÃO

BRUNO ALMEIDA ODIERNA

**IDENTIFICAÇÃO E VALIDAÇÃO DOS TIPOS DE JOGADORES POR MEIO DO SEU
GAMEPLAY**

São Paulo
2018

BRUNO ALMEIDA ODIERNA

IDENTIFICAÇÃO E VALIDAÇÃO DOS TIPOS DE JOGADORES POR MEIO DO SEU
GAMEPLAY

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Computação da Universidade Presbiteriana Mackenzie, como requisito para a obtenção do título de Mestre.

ORIENTADOR: Prof. Dr. Ismar Frango Silveira

São Paulo
2018

O24i Odierna, Bruno Almeida
Identificação e validação dos tipos de jogadores por meio do
seu gameplay / Bruno Almeida Odierna. – São Paulo, 2018.
78 f. : il. ; 30 cm.

Dissertação (Mestrado em Engenharia Elétrica e Computação)
– Universidade Presbiteriana Mackenzie, São Paulo. 2018.

Orientador: Ismar Frango Silveira.
Bibliografia: f. 69-73.

1. Classificação de Jogadores. 2. Analítica de Jogos. 3.
Arquétipos de Bartle. 4. Mineração de Dados de Jogos. I.
Título.

CDD 794

Bibliotecária Responsável: Marta Luciane Toyoda – CRB 8/ 8234

BRUNO ALMEIDA ODIERNA

IDENTIFICAÇÃO E VALIDAÇÃO DOS TIPOS DE JOGADORES POR MEIO
DO SEU GAMEPLAY

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Computação da Universidade Presbiteriana Mackenzie, como requisito parcial para a obtenção do título de Mestre em Engenharia da Computação.

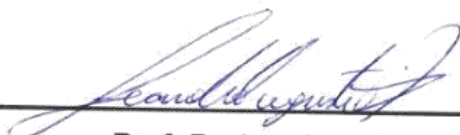
Orientadora: Prof. Dr. Ismar Frango Silveira

Aprovada em 14 de agosto de 2018.

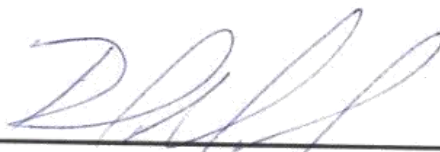
BANCA EXAMINADORA



Prof. Dr. Ismar Frango Silveira
Universidade Presbiteriana Mackenzie



Prof. Dr. Leandro Augusto da Silva
Universidade Presbiteriana Mackenzie



Prof. Dr. Ricardo Nakamura
Escola Politécnica da Universidade de São Paulo

AGRADECIMENTOS

Agradeço a Deus, por me possibilitar fazer esse estudo, e por acolher meus pedidos de iluminação durante as horas mais difíceis.

Ao Prof. Dr. Ismar Frango Silveira, por toda a paciência e direcionamentos durante esse período de orientação, além das palavras positivas e de não me deixar desanimar nos momentos de recusas dos artigos.

Aos meus pais, Isabel e Luiz, pelo incentivo desde sempre na vida acadêmica e na minha formação de caráter.

À minha namorada Marina, pela compreensão, auxílio e apoio.

E aos meus amigos, que sempre entenderam minhas ausências durante o período.

Tudo o que temos de decidir é o que fazer com o tempo que nos é dado

Gandalf

(J.R.R.Tolkien – Senhor dos Anéis).

RESUMO

Analisar e entender o comportamento de jogadores em ambientes virtuais têm sido atividades utilizadas nas empresas desenvolvedoras e produtoras de jogos digitais. Os jogadores não apenas jogam, como também são consumidores de produtos *in game*, compram possíveis sequências ou expansões de jogos, divulgam o jogo para amigos. No caso de jogos *Massive Multiplayer Online Role Playing Game* (MMORPG), os jogadores possuem padrões de jogabilidade variados e, classificando-os, é possível fazer com que os desenvolvedores possam fazer alterações que satisfaçam um maior número de jogadores de modo específico, fazendo com que estes permaneçam em seu ambiente por mais tempo. Esse estudo visa mostrar que é possível classificar jogadores através da análise do *gameplay*, em teorias consolidadas como os arquétipos de Bartle. É apresentada uma análise dos principais tipos de jogadores estudados de formas distintas: análise comportamental, questionários e análise de dados de telemetria de jogos. Também há uma seção dedicada a descrever os processos de *Game Analytics* e uma seção com os resultados da pesquisa.

Palavras-Chave: Classificação de jogadores, Analítica de jogos, Arquétipos de Bartle, Mineração de dados de jogos

ABSTRACT

Analyzing and understanding the behavior of players in virtual environments has been a key activity for companies that develop and produce digital games. Players not only play, but they are also consumers of in-game products, buy possible sequences or expansions of games, and spread them to friends. In the case of Massive Multiplayer Online Role Playing Game (MMORPG), players have different gameplay patterns, and by classifying them, it is possible for developers to make changes that satisfy more types of players in a specific way, resulting in spending more time in the virtual environment. This study aims to show that it is possible to classify players through gameplay analysis in well-established theories such as Bartle's archetypes. It is presented an analysis of the main types of players studied in Academy as follows: behavioral analysis, questionnaires and game telemetry data analysis. There is also a section describing the Game Analytics processes and a section with the results.

Key words: Player classification, Game analytics, Bartle's archetypes, Game data mining,

LISTA DE ILUSTRAÇÕES

Figura 1	Oscilação dos termos de pesquisa <i>Game Metric</i> (azul), <i>Game Analytics</i> (vermelho) e <i>Game Data Mining</i> (amarelo) nos últimos dez anos.....	16
Figura 2	Fluxo de desenvolvimento do projeto de pesquisa.	18
Figura 3	Subdivisão das <i>Player Metrics</i>	25
Figura 4	Os três passos para realizar o clustering, utilizando k-means desde o clustering inicial (a), iteração (b) e o cluster final (c).	33
Figura 5	Projeção para os próximos anos para a indústria de videogames.....	34
Figura 6	Projeção para os próximos anos para a indústria de jogos digitais.	35
Figura 7	Interface do MUD Aardwolf RPG.....	36
Figura 8	Os tipos de jogadores e o impacto no mundo dos jogos r.	37
Figura 9	Resultado do Bartle Test de Andreasen e Downey.	38
Figura 10	Os 4 novos tipos de jogadores.	39
Figura 11	Os oito novos tipos de jogadores de Bartle.	39
Figura 12	Distribuição de Idade por gênero de jogadores (Homens = 4705, Mulheres = 788).....	42
Figura 13	8 tipos de jogadores de Marczewski.....	52
Figura 14	Tipos de usuários de Marczewski.....	54
Figura 15	Número de jogadores ativos no jogo World of Warcraft de janeiro 2005 a setembro 2015, divulgados pela Blizzard.	56
Figura 16	Dados iniciais do WoWAH.	58
Figura 17	Dados adaptados para o experimento.	59
Figura 18	Metodologia usada para identificar os tipos de jogadores, usando k-means.....	61

Figura 19	Framework de <i>Game Analytics</i>	62
Figura 20	Framework de <i>Game Data Mining</i>	Erro! Indicador não definido.
Figura 21	Número de jogadores em cada uma das top 10 <i>guilds</i>	64
Figura 22	Distribuição de Classes de Avatares na <i>Guild 103</i>	64
Figura 23	Mapa de calor das classes e raças dos Avatares da <i>Guild 103</i>	65
Figura 24	Usuários ativos por mês.....	66
Figura 25	Distribuição de horas jogadas de cada jogador da <i>guild 103</i>	66
Figura 26	Horário de <i>Login</i> dos jogadores da <i>guild 103</i> no primeiro semestre de 2008.	67
Figura 27	Horário de <i>Login</i> dos jogadores da <i>guild 103</i> no segundo semestre de 2008.	67
Figura 28	Tempo para atingir o nível máximo de cada classe.....	68
Figura 29	Zonas mais acessadas entre jogadores de cada nível.....	69
Figura 30	Deslocamento global dos jogadores de nível entre 70 e 80 da <i>guild 103</i>	70
Figura 31	Deslocamento nas regiões Arena.....	70
Figura 32	Deslocamento dos jogadores na região <i>Battlegrounds</i>	71
Figura 33	Deslocamentos dos jogadores a partir das Capitais.....	71
Figura 34	Grupos de jogadores a partir da quantidade de horas jogadas.....	72
Figura 35	Quantidade de horas jogadas em cada uma das regiões.	73
Figura 36	Agrupamento de jogadores das regiões Northrend e Outland pela quantidade de horas jogadas e quantidade de visitas nessas regiões.	73
Figura 37	Grupos de todos os jogadores.....	74
Figura 38	Jogadores agrupados pela quantidade de horas jogadas em cada região.....	75

Figura 39 Quantidade de vezes que as regiões Northrend e Outland foram visitadas e quantidade de horas jogadas e quantidade de visitas nessas regiões. 76

LISTA DE TABELAS E QUADROS

Tabela 1	Distribuição de horas jogadas por semana, por gênero.	42
Tabela 2	Características de jogadores e seus agrupamentos.	43
Tabela 3	Distribuição de jogadores por <i>cluster</i>	76
Tabela 4	Distribuição de jogadores por arquétipo.	77
Quadro 1	Grupos, percentual e as características dos jogadores identificados no Tera, por meio do algoritmo k-means.	49
Quadro 2	Grupos, percentual e as características dos jogadores identificados no Tera, através do algoritmo SVIM.	50
Quadro 3	Grupos, percentual e as características dos jogadores identificados no BF2BC2, através do algoritmo k-means.	51
Quadro 4	Grupos, percentual e as características dos jogadores identificados no BF2BC2 através do algoritmo SIVM.	51
Quadro 5	Dados iniciais disponíveis na extração reduzida do WoWAH.	57
Quadro 6	Dados do WoWAH ajustados para o experimento.	58

LISTA DE ABREVIATURAS

ARPU	Average Revenue Per User
B.I.	Business Intelligence
BF2sBC2	Battlefield 2: Bad Company 2
CRISP-DM	Cross-Indutor Standard Process for Data Mining
DSI	Days Since Install
EDA	Análise Explorativa de Dados
FPS	First Person Shooter
GUR	Game User Research
KDD	Knowledge discovery from data
MMO	Massive Multiplayer Online
MMORPG	Massively Multiplayer Online Role Playing Game
MOBA	Multiplayer Online Battle Arena
MUD	Multi-User Dungeon
NPC	Non-Playable Characters
PAM	Partitioning Around Medoids
PvP	Player-vs-player
RPG	Role Playing Game
SIVM	Simplex Volume Maximization
SOM	Self-Organized Map
T.I.	Tecnologia da Informação
WoW	World of Warcraft
WoWAH	World of Warcraft Avatar History

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivos.....	16
1.2	Justificativa.....	16
1.3	Metodologia.....	17
1.4	Estrutura da Dissertação	19
2	REVISÃO DA LITERATURA	20
2.1	GAME ANALYTICS	20
2.1.1	Definições.....	20
2.2	Telemetria de Jogos	21
2.3	<i>Game Metrics</i>	23
2.4	<i>Game Data Mining</i>	25
2.4.1	<i>Data Mining</i>	26
2.4.2	Utilização dos dados de telemetria de jogos	28
2.5	Análise de <i>Cluster</i>	30
2.5.1	Tipos de agrupamentos básicos.....	31
2.6	Definição do algoritmo	32
3	ARQUÉTIPOS DE JOGADORES	34
3.1	A Taxonomia de Richard Bartle	36
3.2	As motivações dos jogadores.....	41
3.2.1	Motivações dos jogadores de Yee	42
3.3	Uso de dados de Telemetria de jogos para definir tipos de jogadores.....	45
3.3.1	Tipos de jogadores em jogos de plataforma.....	46
3.3.2	Tipos de jogadores em MMORPG e Jogos <i>First Person Shooter</i> (FPS).....	47
3.4	Os tipos de jogadores para ambientes gamificados de Marczewski.....	51
3.5	sobre tipos de jogadores.....	54
4	RESULTADOS E DISCUSSÃO	56
4.1	O Grupo de Dados	56
4.2	Preparação dos dados.....	57
4.3	Apresentação dos resultados do experimento inicial.....	63
4.4	Agrupamento e classificação dos jogadores	71
4.4.1	Agrupamento dos jogadores da guild 103.....	71
4.4.2	Agrupamento com todos os jogadores	73

5	CONCLUSÃO E TRABALHOS FUTUROS	78
	REFERÊNCIAS	81

1 INTRODUÇÃO

Jogos digitais fazem parte da vida de muitas pessoas, desde seu surgimento na década de 1950, com o objetivo de entreter os seus jogadores e obter o retorno investido no desenvolvimento dos seus jogos. Para isso, empresas desenvolvedoras de jogos investem tempo e dinheiro na criação de novos jogos, em equipes de programadores, arte, roteiro, *game design* e diversas áreas relacionadas.

O advento dos jogos massivos *online* e a conseqüente geração de volumes consideráveis de dados a respeito dos jogadores e de sua interação com os jogos e seus elementos – o que se convencionou chamar de *gameplay*, fez surgir o interesse na análise de tais dados, em especial por parte das empresas desenvolvedoras. Buscam, assim, entender os comportamentos dos jogadores no ambiente virtual, seus movimentos, observar padrões de compra, acúmulo de capital ou moeda do jogo, entre outros aspectos. Neste contexto, as informações obtidas são utilizadas para que possíveis erros sejam encontrados, ou que a experiência dos jogadores possa ser aperfeiçoada de forma que seu jogo fidelize seus jogadores e que se tenham jogos longevos, com maiores potenciais de rejogabilidade, como pode ser visto em Dankoff (2014).

A análise das características dos jogadores pode ser feita através dos dados gerados no jogo. Esses dados são chamados de telemetria de jogos ou *Game Metric* e são tratados através de técnicas de *Game Analytics* (a mais utilizada é conhecida como *Game Data Mining*) (DRACHEN, EL-NASR, CANOSSA, 2013).

A informação mais importante para se realizar o *Game Data Mining*, porém, não é o código em si construído para realizar a tarefa de mineração, e sim os dados coletados com as informações adquiridas dos jogadores (ZENN, 2012). Como obter os dados, como tratá-los e como extrair as informações são itens a serem discutidos mais adiante neste trabalho.

A Figura 1 exibe a evolução em relação ao interesse a respeito de tais tópicos: por meio de uma análise baseada na API *Google Trends*¹ nos últimos dez anos, tem-se que a pesquisa do termo *Game Metric* tem diminuído, enquanto *Game Analytics* aumenta desde 2007 e ultrapassou *Game Metric* em novembro de 2010. O interesse por *Game Data Mining* tem se mantido constante, apesar de abaixo dos outros dois termos, o que mostra que é um termo correlato ao *Game Analytics*. Porém, por se tratar de aplicação de técnicas de *Data Mining* e

¹ <https://trends.google.com/>

estar muitas vezes dentro de pesquisas da área de Game Analytics, como em Drachen et al. (2013), não é muito pesquisado.

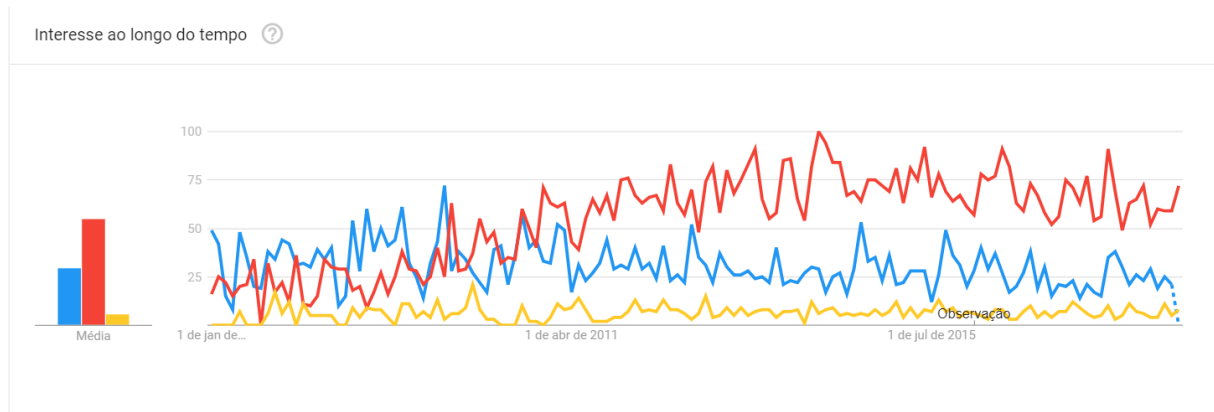


Figura 1 – Oscilação dos termos de pesquisa dos termos *Game Metric* (azul), *Game Analytics* (vermelho) e *Game Data Mining* (amarelo) nos últimos dez anos.

Fonte: Google, 2018

1.1 OBJETIVOS

Este estudo teve os seguintes objetivos:

Principal: analisar dados de jogadores de um jogo do estilo *massively multiplayer online role playing game* (MMORPG) e encontrar os padrões dos mesmos.

Específicos:

1. Organizar e classificar os jogadores em grupos.
2. Analisar sua aderência a propostas de arquétipos bem estabelecidas na literatura.

1.2 JUSTIFICATIVA

A classificação de jogadores é um tema de pesquisa que detém uma razoável atenção da comunidade (tendo 444 resultados encontrados na *IEEEExplore* procurando o termo *Player Classification*) e, dentre as taxonomias, a mais conhecida e utilizada é a de Richard Bartle (1996), sendo citado 2723 vezes, desde 2012, segundo o seu perfil no *Google Scholar*. Por outro lado, não foram encontradas em bases de pesquisas acadêmicas (*Elsevier*, *IEEEExplore* e *Web of Science*), do período de 2009 até o presente momento, trabalhos que relacionem classificação de jogadores, *Game Data Mining*, e taxonomias já conhecidas, uma vez que Bartle (1996) e Marczewski (2015) já possuem seus arquétipos bem definidos através dos seus questionários.

O que foi encontrado nas bases pesquisadas são exemplos de como classificar jogadores por meio do *Gameplay* de jogos de qualquer *Real-Time-Strategy*, como em Drachen et.al

(2009), ou através de questionários, como em Yee (2005). Desta forma, o trabalho se mostra relevante ao relacionar os três tópicos citados a um ambiente *Massive Multiplayer Online Role Playing Game* (MMORPG).

1.3 METODOLOGIA

Inicialmente foi realizada uma revisão da literatura com os termos de busca “*Game data mining*” e “*Game Analytics*” nas bases acadêmicas. Não se tratou de uma revisão sistemática, mas sim uma revisão de literatura para identificar trabalhos correlatos.

A primeira definição a ser feita foi a escolha do MMORPG a ser utilizado como base de dados. Para atingir o objetivo, optou-se pelo estudo dos dados de jogadores do jogo *World of Warcraft*² e foi utilizada a base de dados *World of Warcraft Avatar History* (WoWAH³), pois possui dados reais obtidos durante um período de três anos (2006 a 2009). Ainda que seja uma base de dados relativamente antiga, trata-se de uma base de dados ainda utilizada em algumas pesquisas recentes (BARROS, NOTARGIACOMO 2016; LEE et al. 2011), bem como exercícios de *hackathon* (KAGGLE, 2015).

Em seguida, foram aplicadas técnicas de *Game Data Mining* sobre os dados disponíveis na base de dados WoWAH, para agrupar e classificar os jogadores. Após essa classificação, verificou-se a possibilidade de classificar os jogadores de acordo com uma das taxonomias presentes na revisão de literatura, na qual se destacam as de Bartle (1996) e de Marczewski (2015). Foi escolhida a aplicação do algoritmo k-means, de maneira a gerar *clusters* de jogadores e analisar a relação de tais agrupamentos da taxonomia consolidada na comunidade de pesquisa a respeito de tipos de jogadores: a de Bartle (1996), tendo sido previamente identificada e validada por meio de questionários e observações dos jogadores.

De maneira esquemática, o trabalho seguiu o fluxo da figura 2 a seguir, e foi baseado no trabalho de Drachen, El-Nasr e Canossa (2009), no qual os autores utilizam de uma rede neural SOM (*Self-Organized Map*) para auxiliar na modelagem de jogadores do jogo *Tomb Raider: Underworld*⁴. No caso deste trabalho, a rede SOM foi substituída pelo algoritmo K-means.

² <https://worldofwarcraft.com/>

³ <http://mmnet.iis.sinica.edu.tw/dl/wowah/>

⁴ <https://www.tombraider.com/en-us>

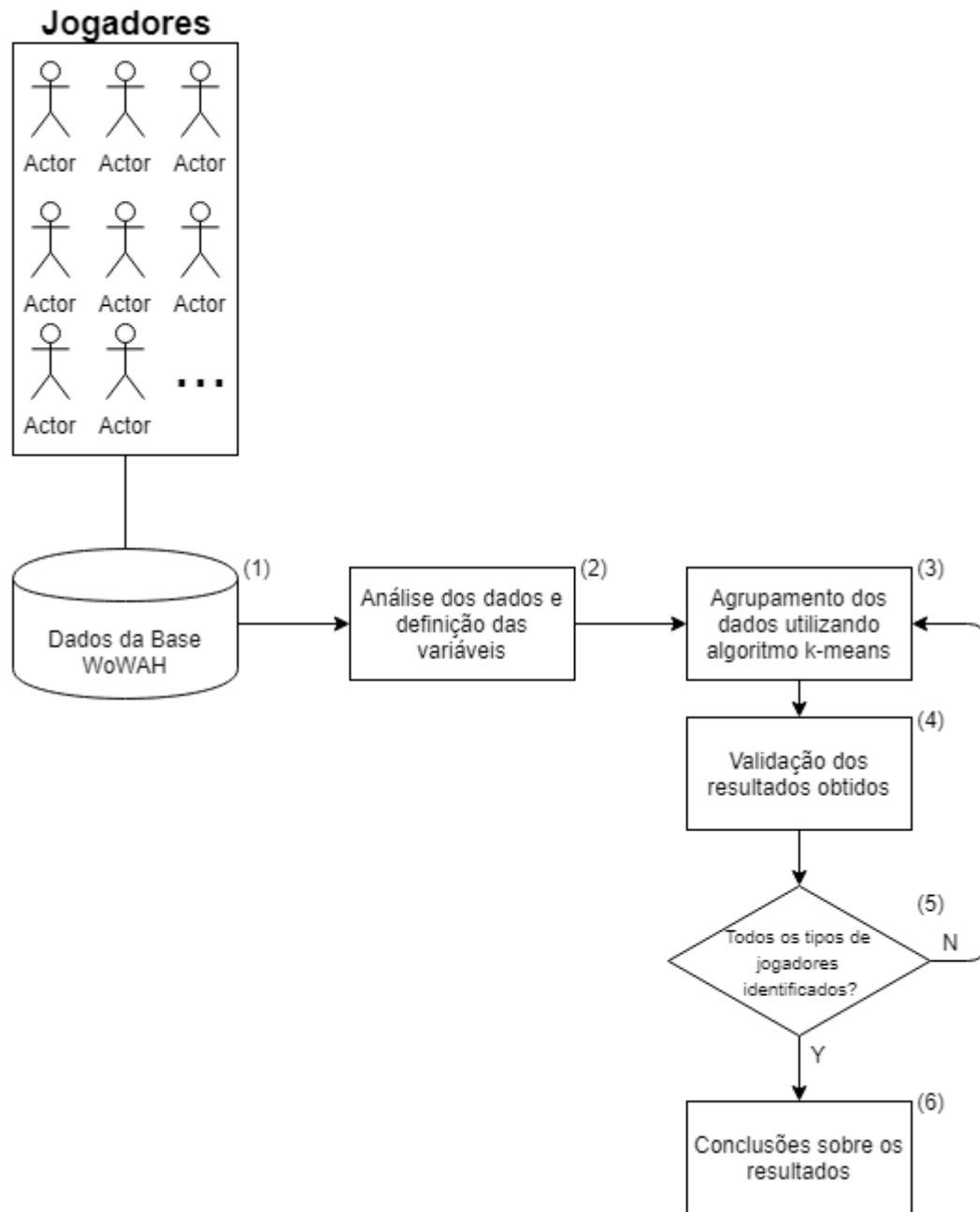


Figura 2 – Fluxo de desenvolvimento do projeto de pesquisa
 Fonte: Criação do Autor

Inicialmente realizou-se o armazenamento da base de dados no banco de dados (1) e posteriormente a definição dos dados que foram utilizados para classificar os jogadores (2). A seguir, os dados foram agrupados pelo algoritmo k-means (3) para separar os valores similares, possibilitando encontrar padrões dos jogadores nos dados armazenados. Neste passo, foram encontrados os grupos de jogadores que melhor se encaixam nas classificações definidas pelas taxonomias adotadas e, durante esse passo foi definido que a taxonomia adotada seria: Bartle (1996), por causa da sua adoção nas pesquisas, e por se tratar de um jogo MMORPG para o qual fora criada. O próximo passo (4) foi verificar a possibilidade ou não de encontrar as

características descritas na taxonomia definida no passo anterior, com o auxílio de análise do *gameplay* de jogadores e do algoritmo k-means. Em caso negativo, seria necessário voltar ao passo (3) e realizar um novo agrupamento dos dados com k-means. Em caso positivo, passaríamos para o próximo passo: relatar as conclusões sobre os resultados obtidos (6).

1.4 ESTRUTURA DA DISSERTAÇÃO

Assim sendo, esta dissertação tem a seguinte estrutura: além desta introdução, no capítulo 1, foram organizados mais quatro capítulos. O capítulo 2 fundamenta o *Game Analytics* e suas possíveis aplicações. O capítulo 3 traz os arquétipos dos jogadores e as diferentes e mais populares abordagens em classificação de jogadores. O conteúdo do capítulo 4 traz os comentários dos testes realizados e seus resultados. Por fim, o capítulo 5 apresenta as considerações finais e indicações de trabalhos futuros.

2 REVISÃO DA LITERATURA

2.1 GAME ANALYTICS

2.1.1 Definições

Como visto anteriormente, a indústria de jogos tem crescido constantemente durante os últimos anos, o que motiva cada vez mais as empresas desenvolvedoras de jogos a investir em novos jogos, sejam eles realmente novos ou continuações de franquias já consolidadas.

Se considerarmos os jogos lançados no ano de 2017, até o mês de novembro foram 334 jogos distribuídos entre as maiores plataformas de jogos (PC, Playstation, Microsoft, Nintendo) (ESPINELI; PLAGGE, 2017) e, se forem levados em consideração os jogos para aparelhos celulares, foram encontrados 3.300.000 jogos disponíveis para download (STATIA, 2017).

Essa grande quantidade de jogos digitais disponíveis compete entre si pela atenção e tempo dos jogadores, e as desenvolvedoras querem que seus jogos sejam jogados e que gerem lucro. Para isso, no processo de desenvolvimento de um jogo, utilizam de métodos e ferramentas para que os jogos sejam mais efetivos em seus objetivos. Um desses métodos é o *Data Analytics* (DRACHEN; EL-NASR; CANOSSA, 2013).

Segundo Drachen, El-Nasr, Canossa (2013), o termo *Analytics* (ou “Analítica”, como se costuma traduzir para o português) refere-se ao processo de descobrir e comunicar padrões em dados, a fim de resolver problemas nos negócios ou realizar previsões para sustentar decisões gerenciais, realizar ações e/ou melhorar desempenho. Nesse sentido, a fundamentação metodológica do *Data Analytics* é a estatística, enquanto *Data Mining* refere-se aos campos da Matemática, programação e inteligência artificial.

Data Analytics não pode ser confundido com análise de dados, uma vez que é um conceito guarda-chuva, que abrange uma metodologia sobre como encontrar e comunicar padrões de dados, enquanto o segundo termo refere-se a analisar algo específico em um *dataset* (DRACHEN; EL-NASR; CANOSSA, 2013) (HAN; KAMBER; PEI, 2012).

Uma ramificação de *Analytics* é o *Game Analytics*, que é a aplicação das ferramentas de *Analytics* aplicadas em desenvolvimento e pesquisas de jogos.

O benefício de se utilizar *Game Analytics* é poder ter um melhor embasamento na hora de tomar alguma decisão relacionada ao seu jogo, desde o *game design* até em ações de marketing, e sempre com um objetivo: levar melhor experiência ao usuário final (DRACHEN; EL-NASR; CANOSSA, 2013).

Um erro comum na área de *Game Analytics* é pensar que só é aplicável para um tipo de análise sobre comportamento de jogadores, ou sobre a fase de testes do jogo. Pode-se, a princípio, usar *Game Analytics* para responder perguntas de diferentes tipos. Mas, apesar de ter um conceito fundamental e ter um grande potencial de usabilidade, *Game Analytics*, diferente de outras subdivisões de *Analytics*, ainda não possui um padrão de termos técnicos e processos definidos especificamente para ele (DRACHEN; EL-NASR; CANOSSA, 2013) (YANNAKAKIS; TOGELIUS, 2015).

Uma área que possui uma grande quantidade de pesquisas realizadas é a de *Game User Research* (GUR), inclusive o presente trabalho se dá nessa área, e utiliza de algumas técnicas e metodologias que advêm da Psicologia, Inteligência Computacional, *Machine Learning*, Interação Homem-Máquina, para avaliar como as pessoas (ou jogadores) interagem e qual a qualidade dessa interação com os jogos (DRACHEN; EL-NASR; CANOSSA, 2013).

2.2 TELEMETRIA DE JOGOS

Para utilizar o processo de descobrimento de padrões em dados, é necessário ter esses dados. Dados são criados a partir do momento em que se inicia uma sessão de jogo virtual até o momento em que este é desligado, e esses dados estão disponíveis para os desenvolvedores e produtoras poderem analisar seus jogadores. Esses dados criados e coletados pelas desenvolvedoras são conhecidos como telemetria de dados de jogos.

A telemetria é uma forma de coleta remota de dados, onde se instala uma aplicação de coleta de dados de interação jogo-jogador do jogo desejado, e as envia digitalmente para uma base de dados, que armazena de forma acessível todas as informações para que seja possível realizar quaisquer tipos de análises do que foi coletado. Esses dados podem ser de qualquer natureza que o jogo permitir (como *timestamps*, localização dos jogadores, pontuação, locais de morte, entre outros que o jogo possa enviar) e podem conter informações do comportamento do jogador dentro do jogo, das transações realizadas, e das conversas com personagens (DRACHEN; EL-NASR; CANOSSA, 2013).

As desenvolvedoras e produtoras utilizam esses dados para formar modelos estatísticos a fim de obter um melhor entendimento sobre seus jogadores e jogos. As dificuldades encontradas são: saber quais dados serão analisados e utilizados para encontrar os modelos desejados e onde serão armazenados os dados colhidos, pois é um volume muito grande de dados e pode-se demorar ao tentar analisar de modo convencional. (BARROS; NOTARGIACOMO, 2016) (DRACHEN; EL-NASR; CANOSSA, 2013).

Como os dados de telemetria são coletados remotamente, coleta-se através de aplicações que são instaladas nos jogos e enviam essas informações para um servidor que armazenará os dados de interação com o jogo (DRACHEN et al., 2009).

Por não existir um padrão único bem definido, os dados transmitidos possuem nomes diferentes dependendo da literatura escolhida. Porém Drachen, El-Nasr e Canossa (2013) definem que a essência da telemetria são os atributos dos objetos, ou itens, que são enviados como o posicionamento do personagem em algum lugar do mapa, a quantidade de moeda corrente do jogo que o jogador possui, quantidade de amigos no jogo, tempo necessário para terminar uma determinada fase, entre outros.

Para trabalhar com os dados obtidos, é necessário transformar o dado bruto em variáveis ou características (dependendo da literatura encontra-se os dois termos que apontam para um mesmo termo). Essas variáveis ou características possuem um domínio específico que são um grupo de todos os valores possíveis para esses atributos e que são determinantes para que seja possível operacionalizar os dados recebidos (DRACHEN; EL-NASR; CANOSSA, 2013).

Os dados de telemetria podem ser enviados de três formas fundamentais: Eventos, Frequência e Ativação, e podem ser enviados para a base de dados quando um certo evento ocorre no jogo; essa forma é conhecida como transmissão por eventos (DRACHEN; EL-NASR; CANOSSA, 2013).

A telemetria pode ser enviada com uma frequência específica, determinada pelos desenvolvedores e analistas, ao invés de ser acionada por um evento específico. Nesse caso, os dados podem ser enviados a cada determinado período de tempo, e será armazenado no servidor (DRACHEN; EL-NASR; CANOSSA, 2013).

Deve ser possível também para a equipe que analisa os dados, ativar e desativar o monitoramento de alguns dados, principalmente para monitorar os dados de jogadores, em novas áreas disponíveis para os jogadores em novas atualizações. Para isso, essa possibilidade de ligar ou desligar o envio de dados é importante para a equipe de desenvolvimento (DRACHEN; EL-NASR; CANOSSA, 2013).

A questão primordial para telemetria de dados é a confidencialidade das informações contidas nos grupos de dados colhidos dos jogadores. Esses dados podem conter informações como nome, idade, localização, informações de cartão de crédito, entre muitas outras que os jogadores não gostariam que fossem divulgadas para terceiros. Por isso, todos os dados são mantidos de forma confidencial e não são vendidos ou compartilhados com nenhuma outra empresa. Esse é um dos motivos pelo qual não existem muitos estudos na academia, com dados reais de jogos desenvolvidos por grandes *softhouses*. Apesar de ser uma área ainda a ser

desenvolvida, está disponível o *Web Analyst Code of Ethics*⁵, que visa mostrar as melhores práticas para tratar com ética os dados recebidos através da telemetria (DRACHEN; EL-NASR; CANOSSA, 2013).

Uma vez que os dados de telemetria estão armazenados em uma base de dados, pode-se transformar em medidas interpretativas, como quantidade de erros encontrados no desenvolvimento, tempo para terminar um determinado *checkpoint*, quantidade de jogadores ativos, *Average Revenue Per User* (ARPU). Essas ou quaisquer outras medidas relacionadas aos jogos digitais são chamadas de *Game Metrics* (métricas de jogos) (DRACHEN, 2012).

2.3 GAME METRICS

As *Game Metrics* são medidas relacionadas a jogos digitais, que podem ser interpretadas de alguma forma, e a vantagem de se usar essas métricas é que possuem as mesmas vantagens que sistemas de *Business Intelligence* (B.I.), como ajudar em tomadas de decisão com relação ao jogo em desenvolvimento. Essas métricas podem ser variáveis ou características, valores calculados entre os dados coletados, ou mesmo qualquer outra informação armazenada (DRACHEN 2012).

Como citado na sessão anterior, as métricas podem ser de qualquer natureza, desde que estejam envolvidas com jogos. Em jogos de tiro em primeira pessoa, pode ser medido a taxa *Kill/Death*; em jogos de corrida, o que pode ser medido é o tempo para completar uma volta no circuito; jogos MMORPG, a quantidade de ouro pode ser uma métrica a ser escolhida para calcular a quantidade de ouro gasto por item, como outras variáveis/características que são conhecidas como *Game Metrics*. Porém, como não há um padrão nos termos utilizados em *Game Analytics*, os nomes dados às métricas possuem variações entre jogos e autores (DRACHEN, 2012) (DRACHEN; EL-NASR; CANOSSA, 2013).

Métricas são calculadas em função de alguma variável, e geralmente essa variável é o tempo. Toda métrica está ligada à alguma medida de tempo, por isso todos os dados analisados e coletados são com relação ao período que já passou, e assim é possível fazer análises preditivas, com base no histórico de comportamento dos jogadores (DRACHEN, 2012) (DRACHEN; EL-NASR; CANOSSA, 2013).

⁵ <http://www.digitalanalyticsassociation.org/?page=codeofethics>

Melon (2009) categoriza as métricas em 3 tipos e Drachen, El-Nasr e Canossa (2013) adaptam essas métricas como *User Metrics* (métricas de usuários), *Performance Metrics* (métricas de desempenho) e *Process Metrics* (métricas de processos).

User Metrics estão relacionadas às pessoas que jogam os jogos, e são vistas tanto quanto jogadores como consumidores. É nessa categoria de métricas que se encontram a maioria de estudos na área de Game Analytics (MELON, 2009) (DRACHEN; EL-NASR; CANOSSA, 2013).

Tudo que é relacionado ao desempenho dos ambientes, de infraestrutura necessária para dar suporte ao jogo, é medido nas *Performance Metrics*. Essas medidas são utilizadas nas fases de garantia de qualidade, para medir o desempenho dos ambientes após as implantações de novos *updates*, ou pequenos *patches* para correção de erros, ou ainda qualquer alteração nos ambientes dos jogos. Essa é uma das áreas mais maduras do *Game Analytics*, por utilizar métodos derivados da área de garantia da qualidade de T.I (tecnologia da informação) (MELON, 2009) (DRACHEN; EL-NASR; CANOSSA, 2013).

Métricas de processos estão relacionadas ao desenvolvimento de jogos em si. Essas métricas estão relacionadas a medir as fases de desenvolvimento do jogo, assim como em processos de gestão de projetos, pois muitas vezes os jogos são desenvolvidos utilizando métodos ágeis, assim como em muitos outros projetos na área de T.I. (tecnologia da informação), o que torna essa métrica também mais consistente na área de *Game Analytics* (MELON, 2009) (DRACHEN; EL-NASR; CANOSSA, 2013).

O Jogador é o motivo principal pelo qual os jogos são desenvolvidos, principalmente voltado a proporcionar melhor experiência aos jogadores, por isso o *User Metrics* é tão importante para o *Game Analytics*. Para entregar um melhor desenvolvimento, Drachen, El-Nasr e Canossa (2013) subdividem essa categoria em 3 tipos: *Customer Metrics*, *Community Metrics* e *Gameplay Metrics*. *Gameplay metrics*, que por sua vez, é subdividida em outras três categorias: *In-game*, *Interface* e *System*. O diagrama pode ser visto na figura 3.

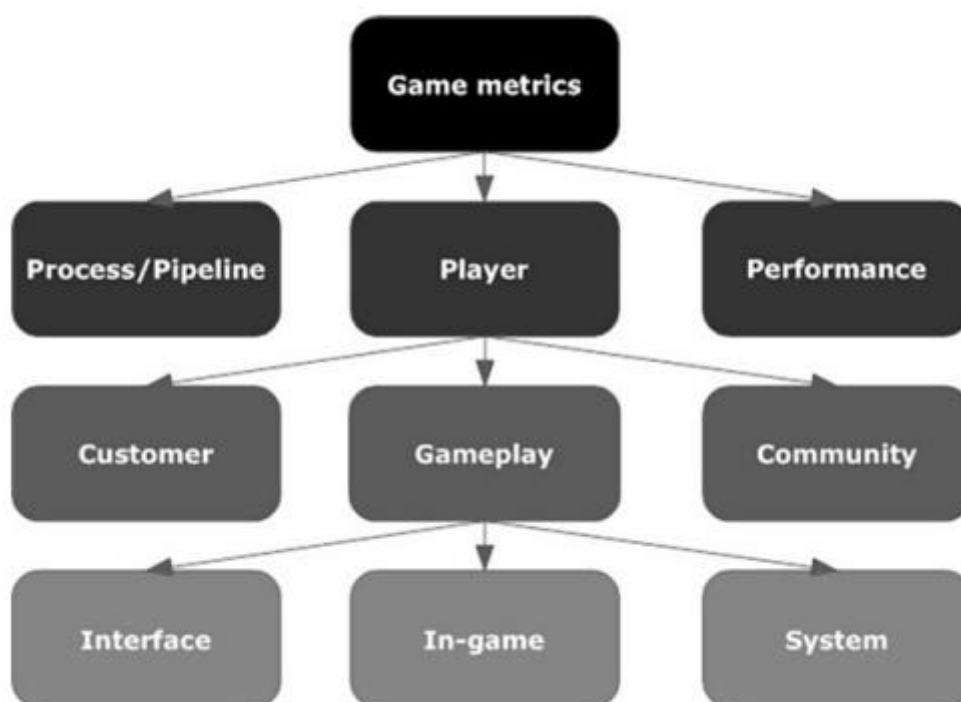


Figura 3 – Subdivisão das *Player Metrics*
 Fonte: Drachen, El-Nasr, Canossa (2013)

Essa divisão mostra como as métricas de usuários são importantes para cada vez mais proporcionar uma melhor experiência para os jogadores, e também importantes para desenvolvedores na hora de criar o jogo. Desenvolvedores analisam as informações dos jogadores e tentam transformar em novas funcionalidades ou atualizações para os jogos (DRACHEN; EL-NASR; CANOSSA, 2013).

2.4 GAME DATA MINING

Uma grande quantidade de dados é gerada em qualquer sessão em ambientes digitais, seja navegando na internet, em transferência de dados através de satélites, ou mesmo em sessões de jogos digitais. Como dados podem ser de qualquer natureza, diferentes áreas têm se aproveitado para realizar pesquisas, como as áreas de biológicas, análise de redes sociais, análise de jogadores, entre outras (DRACHEN et al., 2009) (BARROS; NOTARGIACOMO, 2016).

Esses dados coletados contêm todo tipo de informação sobre seus jogadores, e também monitoram seu próprio ambiente para poder atuar em qualquer anomalia do sistema (DRACHEN et al., 2013).

Apesar de ter acesso a esses dados, nem sempre é uma tarefa fácil a descoberta de informações úteis, devido ao volume. Uma das dificuldades que podem ser encontradas é trabalhar com dados em larga escala, pois não é apenas uma tarefa de recuperar dados de uma base, mas sim transformar esses dados em informações úteis para tomar as ações necessárias. Por isso, foram desenvolvidos métodos que pudessem ajudar a trabalhar com esses dados de larga escala – da ordem de GB ou maiores, – que foram chamados de *data mining* (DRACHEN et al., 2013).

2.4.1 *Data Mining*

Data mining é um termo que não possui uma definição geral única. Segundo Han, Kamber e Pei (2012), *data mining* é um processo que busca descobrir padrões e conhecimentos em grande quantidade de dados. O termo em questão faz referência às antigas técnicas de mineração de ouro nas grandes minas, uma vez que havia de se tirar uma quantidade muito grande de minério para conseguir uma pequena gema de ouro.

Alguns autores tratam *data mining* como descoberta de conhecimento através de dados *Knowledge discovery from data* (KDD). Ainda poderíamos nos referir à *data mining* também como mineração de conhecimento através de dados.

Outros autores ainda consideram *data mining* como um passo essencial do processo de extrair conhecimento dos dados, que segue a sequência a seguir: Limpeza dos dados (remover as inconsistências e os ruídos dos dados); Integração (combinação de uma ou várias base de dados); Seleção (seleção dos dados necessários para a análise); Transformação (dados são transformados e consolidados no formato para realizar a análise); Mineração (processo onde são utilizados métodos inteligentes para a extração do conhecimento); Avaliação dos padrões (investigação dos dados descobertos); e, por fim, Apresentação (utilização de técnicas de visualização de dados para apresentar o conhecimento extraído). (HAN; KAMBER; PEI, 2012).

Não se pode confundir *data mining* com estatística, uma vez que este utiliza muitas ferramentas da estatística para analisar grande quantidade de dados. A estatística seria um grupo de métodos para serem utilizados num grande processo denominado *data mining* (DRACHEN et al., 2013) (HAN; KAMBER; PEI, 2012).

Apesar de o termo *data mining* estar muitas vezes atrelado ao conceito de automatização, ou por utilizar métodos semiautomáticos, a presença de pessoas é fundamental para o sucesso dos processos de *data mining*. São pessoas que desenvolvem os jogos, criam as histórias, e, principalmente, os jogam (DRACHEN et al., 2013).

Para se fazer o *data mining*, foi criado um processo com as melhores práticas utilizadas por empresas e desenvolvedores chamado *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Esse processo pode ser descrito em seis fases distintas: Entender o negócio, entender os dados, Preparação dos dados, Modelagem, Avaliação e Implementação. Como todo processo padrão, ou metodologia, pode ser complicada a implementação de todos os passos. Nesses casos é importante entender, e utilizar da melhor maneira, cada passo do processo para responder as questões a serem descobertas nos dados de telemetria. Cada passo é facilmente adaptável e não precisa ser realizado em sequência, o próximo passo sempre é relacionado à saída do processo anterior. A vantagem de se utilizar esse processo é que é de fácil adaptação aos modos de desenvolvimento ágeis, que são muito comuns na área de desenvolvimento de jogos (MELLON, 2009) (SHEARER, 2000) (DRACHEN et al., 2013).

As seis fases do CRISP estão descritas a seguir: O primeiro processo é entender o negócio, que nada mais é que definir quais são os objetivos e como atingi-los, e modelá-los de forma que seja possível aplicar o *data mining* (SHEARER, 2000) (DRACHEN et al., 2013).

No processo de compreensão os dados, é realizada a coleta dos dados da base, obtém-se as primeiras ideias do que se tem de informações úteis, e avalia-se a qualidade dos dados obtidos (SHEARER, 2000) (DRACHEN et al., 2013). A seguir, inicia-se a preparação dos dados que serão escolhidos para serem utilizados nos passos seguintes. São selecionadas as variáveis desejadas para as análises, removem-se dados ou ruídos indesejados, e consolidam ou transformam variáveis, a fim de obter uma nova variável que será utilizada nas análises (SHEARER, 2000) (DRACHEN et al., 2013).

Seleciona-se e aplica-se as técnicas de *data mining* no processo de modelagem. Esse processo pode se repetir quantas vezes for necessário, e diferentes técnicas podem ser utilizadas para solucionar o mesmo problema (SHEARER, 2000) (DRACHEN et al., 2013).

Avaliar os resultados, medir o desempenho, verificar se todos os objetivos foram satisfatórios, e as conclusões sobre os resultados são realizadas no processo de avaliação (SHEARER, 2000) (DRACHEN et al., 2013).

Por fim, na fase de implementação os resultados são apresentados para as partes interessadas, de modo que sejam de fácil entendimento para o seu público alvo. Podem ser apresentados de formas variadas, dependendo do tipo de audiência (SHEARER, 2000) (DRACHEN et al., 2013).

2.4.2 Utilização dos dados de telemetria de jogos

O primeiro passo para utilizar os dados armazenados é fazer uma exploração nas informações obtidas. Independentemente do formato de armazenamento dos dados, há três formas fundamentais para explorá-los: exploração detalhada ou vertical, generalizada ou horizontal, e específica em um período de tempo. Apesar de serem consideradas básicas, muitas vezes apenas realizando essas explorações pode-se obter os resultados desejados sem precisar de outros métodos mais sofisticados (DRACHEN et al., 2013).

Como os nomes sugerem, cada uma destas formas é utilizada para olhar para os dados de forma específica. Na exploração específica no tempo, como o próprio nome sugere, explora-se os dados em um período de tempo e verifica-se as variáveis no mesmo (DRACHEN et al., 2013).

Na exploração generalizada (ou horizontal), múltiplas variáveis são exploradas ao mesmo tempo, para obter informações dos dados. Nessa verificação, uma variável complementa a informação da outra, podendo ser utilizada para fazer rankings de jogadores que mais gastam horas no jogo, quantidade de dinheiro gasta em uma determinada região, quantidade de jogadores em *guilds*, ou contagem de amigos por jogador (DRACHEN et al., 2013).

Quando se deseja saber detalhes sobre o comportamento dos jogadores, utiliza-se a exploração detalhada (ou vertical). Essa exploração ajuda a encontrar padrões que não são encontrados nas camadas superiores dos dados, como padrões de esquiva de inimigos e movimentação com rotas não comuns (DRACHEN et al., 2013).

Utilizando essas explorações é possível encontrar padrões nos dados, e medir e avaliar esses padrões é fundamental para o processo de *data mining* em geral. Com essas medidas é possível eliminar dados ou padrões indesejados da sua base, classificar os padrões de acordo com a relevância para a pesquisa e utilizar os melhores padrões para realizar a apresentação dos dados (LAROSE, 2004).

2.4.2.1 Métodos de *data mining*

Dentre os dados analisados, é possível realizar alguns tipos de informações que podem ser mineiradas, sendo os mais conhecidos: Caracterização e Discriminação, *Data Mining* de padrões, Associações e Correlações, Classificação e Regressão, Análise de agrupamentos. É possível dividir os diversos métodos em dois duos de categorias: *data mining* descritivo ou preditivo e aprendizado supervisionado ou não supervisionado (HAN; KAMBER; PEI, 2012).

Tentar descrever padrões encontrados em dados do passado, para poder entender como podem influenciar eventos futuros é chamado de *data mining* descritivo (HALOBI, 2017).

No caso de querer prever o que pode acontecer com base nos dados obtidos, utiliza-se o *data mining* preditivo. É sempre bom atentar que todos os métodos de predição não possuem garantia de acerto absoluto, e sim são métodos para encontrar a melhor possibilidade de ocorrer, com base nos dados da telemetria (HALOBI, 2017).

Aprendizado supervisionado é quando um algoritmo é treinado com um conjunto de dados que possuem as características desejadas a fim de chegar à resposta esperada. Após o treinamento é esperado que, quando o algoritmo for apresentado aos dados, seja capaz de informar a saída, caso o valor de entrada seja válido. No caso de jogos, esse método é muito utilizado para saber quando jogadores irão parar de jogar, ou saber quais tipos de itens jogadores irão comprar. Aprendizado supervisionado pode ser associado à classificação dos resultados (REVISTABW, 2015) (DRACHEN et al.,2013) (HAN; KAMBER; PEI, 2012).

Quando o algoritmo aprende com o conjunto de dados sem a necessidade de conhecer a saída desejada, é chamado aprendizado não supervisionado. Nesse caso, não é esperado prever valores de saída, mas sim encontrar relações entre os dados analisados, que é o que na maioria das vezes espera-se encontrar ao utilizar os algoritmos de *data mining*. O aprendizado não supervisionado é sinônimo de agrupamento de informações (MACIEL, 2008) (DRACHEN et al.,2013) (HAN; KAMBER; PEI, 2012).

No contexto de *game data mining*, há métodos que são os mais popularmente utilizados: Descrição, Caracterização, Discriminação, Classificação, Estimativas, Predição, Agrupamento, Associação por afinidade.

Descrição é quando tenta-se descrever as classes individuais e encontrar tendências nos dados dos jogos. A forma mais comum é aplicando *Explorative Data Analysis*, ou Análise Explorativa de Dados, (EDA), que é uma forma gráfica de explorar dados e encontrar padrões ou tendências nos mesmos. Pode-se utilizar esse método para encontrar relações entre quantidade de horas jogadas e os níveis de cada personagem (DRACHEN et al.,2013) (HAN; KAMBER; PEI, 2012).

Para encontrar uma regra característica de uma determinada variável, utiliza-se a Caracterização. Um caso comum é caracterizar os jogadores que completam um determinado checkpoint de uma fase em menos do que 5 minutos, ou quantos jogadores terminaram a fase 3 de um determinado jogo (DRACHEN et al.,2013) (HAN; KAMBER; PEI, 2012).

A Discriminação possui a mesma finalidade da Caracterização, porém o resultado é uma comparação entre duas ou mais variáveis. Nesse caso, pode-se saber quantos jogadores de

idades entre 15-25 anos e 26-35 realizaram compras dentro do jogo (HAN; KAMBER; PEI, 2012) (DRACHEN et al., 2013).

Classificação é um método muito utilizado no desenvolvimento do jogo, uma vez que é usada para organizar os dados em classes. Com essas classes, os algoritmos podem ser treinados e utilizados para serem aplicados em outros métodos, ou em algum outro grupo de dados.

Similar à Classificação, a Estimação pretende descobrir valores numéricos, e não classes. Estimar quantos jogadores irão evadir do jogo após um período de tempo, ou qual será o lucro ao lançar uma determinada expansão são exemplos de como a Estimação pode ser utilizada.

A ideia de utilizar um grande volume de dados para prever possíveis valores futuros é a base da Predição. É um dos métodos mais utilizados na análise de dados de jogos MMOG, devido à importância de se prever como as mudanças aplicadas poderão impactar os jogadores, ou se alterações no jogo são necessárias, pois isso afeta diretamente os lucros das desenvolvedoras desses jogos.

Apesar de possuir semelhança com a Classificação, o Agrupamento possui a diferença de não possuir os nomes das suas classes. Os grupos são criados a partir do que é descoberto e avaliado pelo algoritmo como válido. O algoritmo de Agrupamento procura dividir os dados em grupos, com base na sua semelhança (DRACHEN et al., 2013).

O objetivo da Associação é encontrar atributos que estejam ligados entre si, e criar regras entre eles. Para que seja válido, é necessário encontrar essa regra, calcular a frequência onde é encontrado o atributo, e a probabilidade de que, encontrando um atributo, o outro será realizado (DRACHEN et al., 2013).

2.5 ANÁLISE DE *CLUSTER*

Como mencionado anteriormente, uma das ferramentas para analisar dados de telemetria de jogos é a análise por agrupamento, que é o processo de agrupar os dados em pequenos grupos (*clusters*). Cada *cluster* agrupa dados similares, que são distintos entre si (HAN; KAMBER; PEI, 2012) (ZHANG et al., 2016).

A análise de *cluster* pode ser utilizada como única ferramenta de mineração de dados para observar como os dados são agrupados e, a partir dos grupos, obter informações específicas das características obtidas. A principal vantagem em se utilizar a análise de *cluster* é que se consegue encontrar automaticamente grupos de dados (HAN; KAMBER; PEI, 2012).

Por ser um ramo da estatística, a análise de *cluster* tem sido muito estudada com um foco maior na análise de clusters baseados na distância dos dados. Análises utilizando *k-means*, *k-medoids* têm sido implementadas em bibliotecas de linguagens de programação como R e Python (HAN; KAMBER; PEI, 2012).

2.5.1 Tipos de agrupamentos básicos

Por haver um grande número de algoritmos na literatura acerca do agrupamento de dados, muitas vezes é difícil categorizá-los em grupos, pois podem possuir características que se sobrepõem em diversas categorias. Porém, é mais didático apresentar os agrupamentos organizados pelas suas principais categorias. Os agrupamentos podem ser divididos em quatro grupos: Particionamento, Hierárquico, Baseados em Densidade e Baseados em *Grids* (HAN; KAMBER; PEI, 2012).

Métodos de particionamento dividem uma quantidade de n dados em k partições, onde $k \leq n$. Isso faz com que os dados sejam divididos em k grupos, onde cada um possui ao menos um objeto. A maior parte dos métodos de particionamento são baseados em distância, onde, dado um k (número de partições desejadas), o método cria grupos iniciais e utiliza uma técnica iterativa de realocação dos dados entre os grupos, até formarem os grupos com dados distintos entre eles. Esse método possui uma limitação computacional, devido à grande quantidade de possibilidades para encontrar todas as possíveis partições. Nesses casos, são utilizados os algoritmos como *k-means* e *k-medoids*, que progressivamente otimizam a qualidade dos agrupamentos e das distâncias (HAN; KAMBER; PEI, 2012).

Os métodos hierárquicos criam uma decomposição hierárquica da base de dados, e pode ser classificado como sendo tanto aglomerativo ou divisivo, baseado na forma em que a decomposição é agrupada. O modo aglomerativo, também conhecido como *bottom-up*, começa com cada objeto, formando um grupo separado, que, a cada iteração, agrupa os objetos próximos até que estes estejam agrupados em um único grupo (o nível mais alto da hierarquia), ou até que atinjam o critério de parada. O modo divisivo, também conhecido como *top-down*, começa com os objetos em um único grupo e, a cada iteração, são divididos em grupos menores, até que estejam cada um em um grupo ou que o critério de parada seja atingido. Esse método possui a limitação de que, uma vez que cada passo (agrupar ou dividir) é realizado, não se pode mais desfazer a iteração, porém essa limitação ajuda na menor quantidade de cálculos computacionais (HAN; KAMBER; PEI, 2012).

A maioria dos métodos de particionamento são baseados na distância entre os objetos, e esses métodos têm dificuldades em encontrar grupos em formas não esféricas. Para isso, foram

desenvolvidos métodos que se baseiam na densidade, onde a intenção é aumentar um determinado *cluster* com base na densidade de objetos contidos nesses grupos, até que atinjam um determinado limite (HAN; KAMBER; PEI, 2012).

Os métodos baseados em grades quantificam o espaço amostral em um número finito de células, que formam uma estrutura de grade. Esse método tem a vantagem de possuir um tempo de processamento rápido, que independe da quantidade de objetos, e depende apenas da quantidade de células no espaço. Utilizar as grades é eficiente para problemas de *data mining* espacial (HAN; KAMBER; PEI, 2012).

2.6 DEFINIÇÃO DO ALGORITMO

Para esse trabalho foi escolhido utilizar o método de agrupamento de particionamento, uma vez que os dados estão distribuídos em uma quantidade conhecida, e que há a necessidade de encontrar um número definido de grupos de jogadores (quatro), que é a quantidade descrita em Bartle (1996). A escolha ficou entre os dois algoritmos mencionados anteriormente: *k-means* e *k-medoids*.

K-means é um algoritmo onde o número de *clusters* (k) é escolhido pelo usuário, e cada centroide é inicializado em posições distintas do grupo de dados. Após a inicialização, os centroides são iterados e, baseado na distância Euclidiana e na média dos *clusters*, os centroides se movem e começam a agrupar os dados em *clusters*, até quando não há mais necessidade de realizar movimentos dos centroides de cada *cluster* e o agrupamento é finalizado. Esse processo é descrito na figura 4, onde a figura (a) tem o grupo inicial dos dados divididos em $k=3$ grupos, e os centroides inicializados aleatoriamente nos grupos; na figura (b) ocorre a iteração e os centroides vão se aproximando dos objetos de valores similares; e em (c) o novo grupo está formado com os objetos agrupados. O pseudocódigo pode ser encontrado abaixo também (HAN; KAMBER; PEI, 2012) (ZHANG et al., 2016).

Pseudocódigo do algoritmo *k-means*:

a) *Input*:

k : número de clusters.

D : grupo de dados que contém n objetos.

b) *Output*: dados agrupados em k grupos

Método:

a) Escolher arbitrariamente o número k objetos de D como centros iniciais de cada grupo;

b) **Repetir**;

- c) (Re)atribuir cada objeto ao grupo que possui maior similaridade, baseado no valor da média de cada um no grupo;
- d) Atualizar as médias do grupo, e calcular a nova média dos objetos de cada grupo;
- e) **Até** sem alteração (HAN; KAMBER; PEI, 2012).

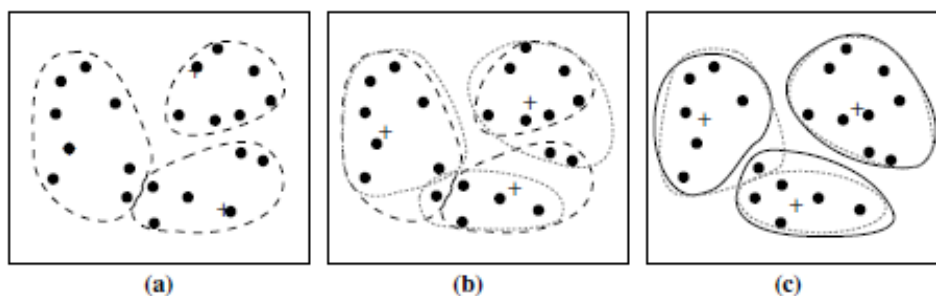


Figura 4 – Os três passos para realizar o *clustering*, utilizando k-means desde o *clustering* inicial (a), iteração (b) e o *cluster* final (c).

Fonte: Han; Kamber; Pei, 2012

O *k-medoid* é similar ao *k-means*, agrupando os objetos através da distância entre eles, porém a sua inicialização se dá em um objeto de fato do grupo de dados. Isso minimiza o erro, uma vez que os pontos extremos do grupo de dados não ficarão à mercê de haver um ponto mais próximo para que sejam agrupados. Uma aplicação com o *k-medoids* é o algoritmo *Partitioning Around Medoids* (particionamento em volta de medoids) (PAM), porém esse algoritmo funciona bem com bases de dados onde o valor de *k* é baixo. Quando o valor é alto, o custo computacional fica cada vez mais alto, tornando esse método pior que o *k-means*.

Apesar do *k-medoid* ser um algoritmo mais robusto, e possuir uma variação no cálculo das médias, o *k-means* ainda possui um custo computacional mais baixo e, na academia, é mais utilizado para a classificação de jogadores como em Drachen et al. (2012). Mesmo fora da literatura de jogos digitais esse método tem sido empregado para agrupar e classificar dados, como em Zhang et al. (2016).

Apesar do *k-means* ser um método efetivo, ainda possui pontos a melhorar, como as informações do grupo de dados que devem ser numéricas, uma vez que se utiliza a média das distâncias entre os centroides de cada *cluster*, além da forma de escolha da quantidade de *clusters* a serem selecionados inicialmente. Apesar de existir algumas formas de melhor escolha através do erro e erro quadrático, para esse trabalho o número de *clusters* (*k*) foi determinado pela quantidade de grupos de jogadores da Taxonomia de Bartle (*k*=4).

3 ARQUÉTIPOS DE JOGADORES

Jogos digitais têm sido uma área de constante crescimento nos últimos 5 ou 6 anos, tendo crescido globalmente, em 2011, de \$54,15 bilhões de dólares para uma projeção, em 2017, de \$78,61 bilhões de dólares, apenas no setor de videogames (jogos para consoles e PC), segundo a Statia (2017). A projeção para os próximos anos está representada na figura 5 abaixo:

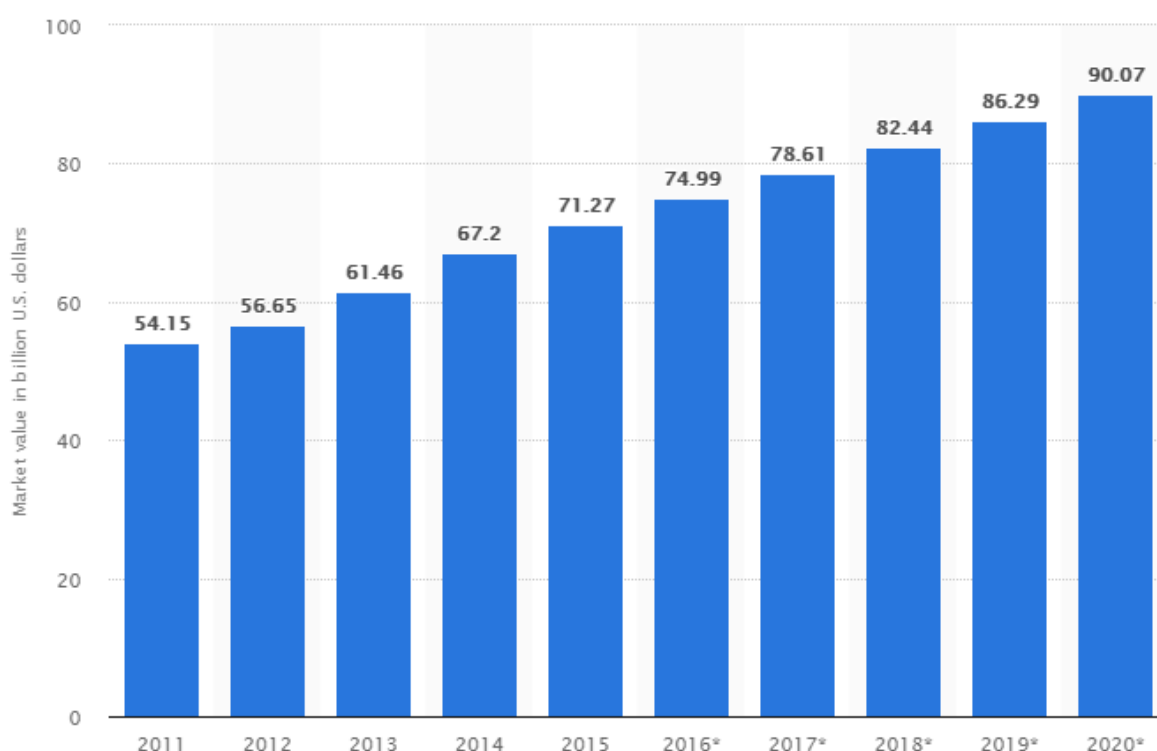


Figura 5 – Projeção para os próximos anos para a indústria de videogames.
Fonte: Statia, 2017

No ano de 2017, para o grupo de jogos como um todo, a estimativa foi de uma movimentação de \$108,9 bilhões de dólares, com jogos de PC e videogames representando 58% desse valor. A projeção é que essa proporção diminua com o advento do crescimento de jogos para *Tablets* e *Smartphones*, que tendem a ocupar 50% em um espaço de três anos (MCDONALD, 2017), como pode ser observado na Figura 6, a seguir.

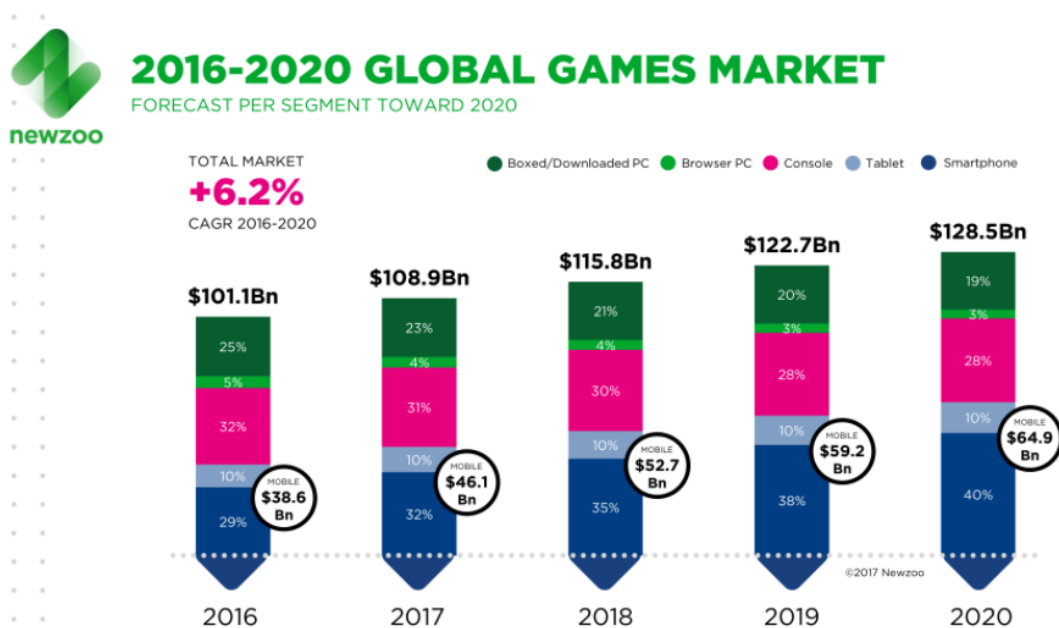


Figura 6 – Projeção para os próximos anos para a indústria de jogos digitais.
Fonte: McDonald, 2017

Com esse volume de dinheiro sendo movimentado dentro da indústria, ficou cada vez mais importante entender o comportamento dos jogadores dentro dos jogos, os gostos, os padrões de jogabilidade, entre outras características que podem determinar o sucesso ou o fracasso de um determinado jogo.

Desde que Richard Bartle (1996) realizou a pesquisa inicial sobre os tipos (ou arquétipos, como vieram a ser conhecidos) de jogadores que jogam MUD (*Multi-User Dungeon*), sua classificação foi aplicada em áreas diferentes da identificação de jogadores de MUD, como Game Design e gamificação (SCHELL, 2008), (SULLIVAN, 2009), (DIXON, 2011), (JORDAN, 2014). Há outros pesquisadores tentando encontrar taxonomias que se ajustem melhor a determinados aspectos dos jogadores, de forma a buscar responder e entender sua motivação em jogar determinados jogos (WORTH; BOOK, 2014), como em Lazzaro (2003), Yee (2005) e Marczewski (2015), por exemplo. Entretanto, para aferir os tipos de jogadores como na concepção original de Bartle (1996), em geral encontram-se apenas questionários, como o desenvolvido por Andreasen e Downey (2001), que valida o tipo de jogador em um MUD ou MMORPG (*Multiplayer Massive Online Role Playing Game*).

A seguir serão mostrados os principais estudos na área de classificação de jogadores, baseados nos estudos desenvolvidos por Bartle (1996), Yee (2006), Drachen et.al (2009), e Marczewski (2015).

3.1 A TAXONOMIA DE RICHARD BARTLE

Bartle (1996) desenvolveu uma taxonomia para categorizar jogadores de MUD (*Multi User Dungeon*). MUDs podem ser considerados um dos primeiros jogos do tipo MMO (*Massive Multiplayer Online*), onde jogadores se juntam para jogar todos em um ambiente virtual, onde todas as interações são realizadas através de textos (BARTLE, 1999). Pode-se ver um exemplo de MUD na figura 7, que possui o mapa da masmorra explorada, as ações realizadas, quantidade de vida, magia, movimentos restantes, a história do mapa explorado, entre outras informações necessárias ao jogador.

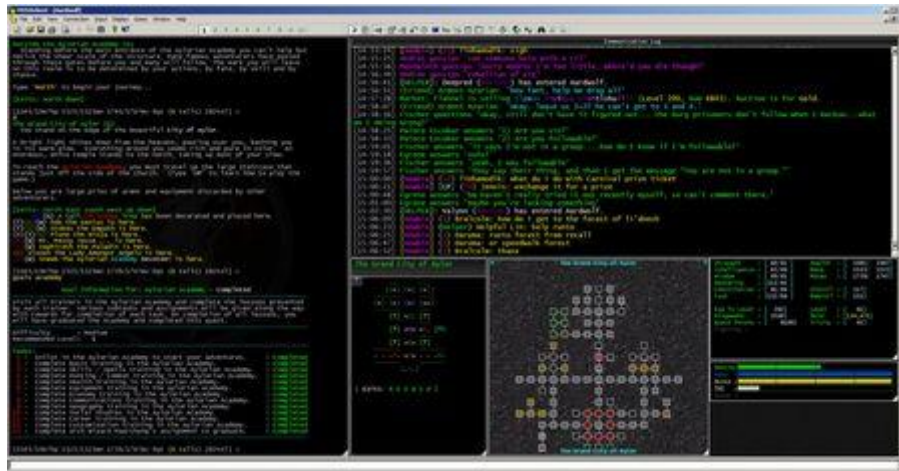


Figura 7 – Interface do MUD Aardwolf RPG⁶.

Fonte: AARDWOLF, 2017⁷

Os tipos de jogadores de Bartle foram resultado de um intenso debate entre Bartle e alguns usuários de um MUD (15 jogadores mais ativos na discussão e outros 15 que não eram tão ativos assim), sobre o porquê de jogarem o jogo, o que mais e o que menos atraía, e o que eles gostariam de mudar ou melhorar (BARTLE, 1996).

Após as respostas se esgotarem, Bartle as organizou e começou a perceber um padrão entre as respostas, dividindo-as em quatro pequenos subgrupos. Apesar de os participantes mostrarem ter um pouco de cada um dos quatro grupos, todos apresentaram uma preferência para cada um dos quatro (BARTLE, 1996).

Os quatro tipos de jogadores foram divididos em: *Killers*, *Explorers*, *Achievers*, *Socializers*, como pode ser visto na figura 8.

⁶ <http://www.aardwolf.com/>

⁷ <http://www.aardwolf.com/images/aardwolf-mud-client4s.jpg>

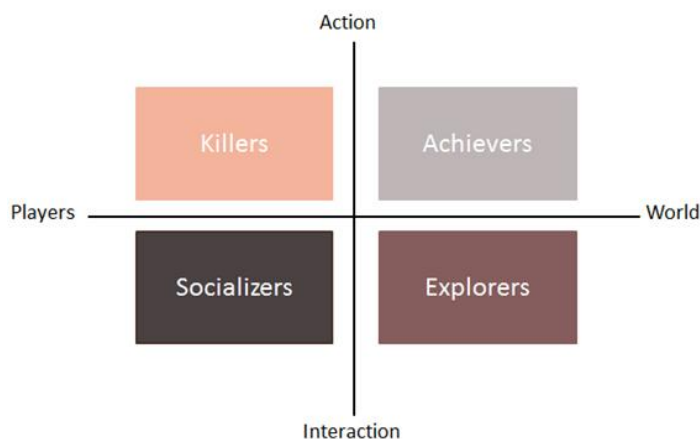


Figura 8 – Os tipos de jogadores e o impacto no mundo do jogo.
 Fonte: Criação do Autor, baseado em (BARTLE, 1996)

O eixo horizontal começa à esquerda na ênfase no Jogador e termina à direita no Ambiente (Mundo do jogo), enquanto o vertical, na parte de baixo começa na interação e termina acima na ação. Os jogadores estão dispostos em cada um dos quadrantes, a partir da definição dos tipos de jogadores acima.

Segundo Bartle (1996), os *Achievers* são os jogadores que jogam motivados por colecionar o maior número de itens possíveis dentro de um jogo, seja uma maior pontuação, objetos raros ou mesmo pequenas recompensas do jogo. A característica desses jogadores é de tentar encontrar alguma vantagem ao participar de alguma missão em grupos, contar a quantidade de experiência que falta para ganhar um nível, ou a quantidade de pontos para se atingir a pontuação máxima. Os *Achievers* gostam de agir sobre o ambiente dos jogos em que se encontram.

Os *Explorers* jogam com a motivação de explorar ao máximo o jogo, não apenas no sentido do mapa do jogo, como também cada detalhe, cada falha, cada abertura no jogo para que se possa tirar alguma vantagem do ambiente. Um Explorer se sente realizado quando encontra algo que ninguém nunca encontrou, seja uma rota escondida ou uma combinação de jogadas nunca tentada, que ativando um poder especial de um personagem. *Explorers* gostam de interagir com o ambiente dos jogos em que se encontram (BARTLE, 1996).

A motivação dos *Socializers* é a comunicação com os outros jogadores, no caso de um jogo de RPG (*Role Playing Game*), representar o seu personagem na melhor forma possível no ambiente virtual. O jogo para esses jogadores não é o principal, e sim poder conversar com os amigos feitos neste ambiente, poder compartilhar experiências fora do jogo, e até mesmo realizar encontros para que todos possam jogar juntos no mesmo local. *Socializers* tendem a uma maior interação com os jogadores do ambiente virtual (BARTLE, 1996).

Os *Killers* são os jogadores que gostam de se impor perante as pessoas do ambiente virtual. São os jogadores que se sentem realizados matando personagens alheios, ou quando provocam outros jogadores durante a sessão de jogo. *Killers* são os jogadores que agem sobre os jogadores que estão no mesmo ambiente que eles (BARTLE, 1996).

O trabalho de Bartle é mencionado em 1830 trabalhos acadêmicos diferentes, segundo o Google Acadêmico em 2017; por ser uma taxonomia de fácil entendimento e usabilidade. Porém, o próprio Bartle (1996) diz que sua taxonomia vem sendo utilizada de forma incorreta, pois ela fora pensada apenas para jogos e mundos do tipo MUD/MMORPG (CASUAL CONNECT EUROPE, 2012).

Há na literatura um estudo desenvolvido por Andreasen e Downey (2001), que classifica os jogadores de MMORPG através de um questionário, que define qual tipo de jogador é mais dominante no usuário que está sendo testado. Esse teste não pode ser dado como certo, uma vez que é apenas uma definição daquele momento do jogador, e depende muito do tipo de jogo que está sendo jogado, mas é uma boa ferramenta para auxiliar na hora de realizar experimentos, indicando perfis para um jogo específico. A figura 9 mostra um resultado do teste de Andreasen e Downey (2001).



Figura 9 – Resultado do Bartle Test de Andreasen e Downey.
Fonte: Andreasen e Downey (2001)

Após a proposição dos 4 tipos de jogadores, Bartle (2005) reafirma que seu modelo não é à prova de falhas. Os dois pontos que tenta corrigir são: seu modelo sugere que jogadores se movem entre as categorias conforme o tempo, porém não demonstra como; todos os quatro tipos de jogadores aparentam ter outros subtipos, mas o modelo inicial não os mostra.

O autor tenta resolver essas falhas adicionando um novo eixo, transformando o modelo antigo de 2 dimensões para um modelo de 3 dimensões. Se antes os eixos eram compostos da ação ou interação dos jogadores com jogadores ou o mundo virtual, esse novo eixo contém se as ações/interações são implícitas ou explícitas.

Ações implícitas são aquelas que são feitas naturalmente, sem consciência, enquanto as ações explícitas são aquelas que foram planejadas previamente (BARTLE, 2005). Assim foram propostos 4 novos tipos de jogadores: *Politicians*, *Networkers*, *Friends*, *Griefers* como visto na figura 10 (BARTLE, 2003).

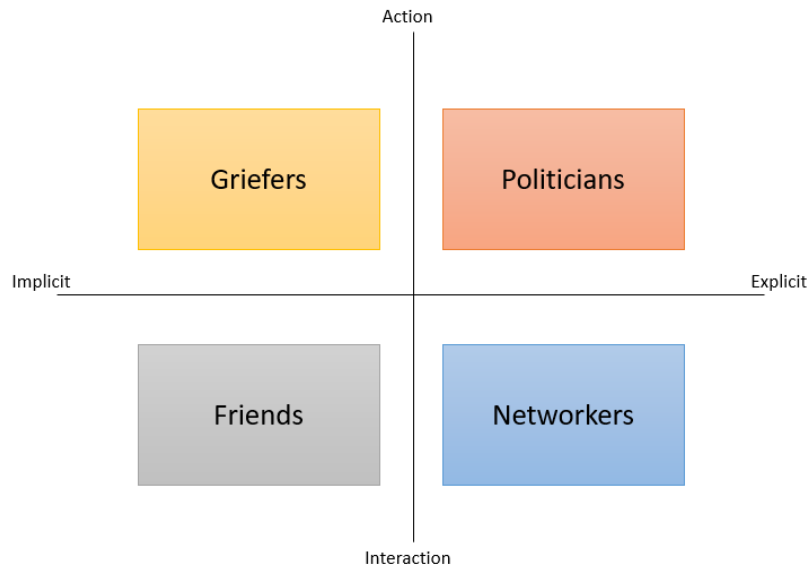


Figura 10 – Os 4 novos tipos de jogadores
Fonte: Criação do Autor, baseado em (BARTLE, 2003)

Unindo os eixos Action/Interaction, Players/World, e Implicit/Explicit e formando um gráfico tridimensional temos a Figura 11 a seguir.

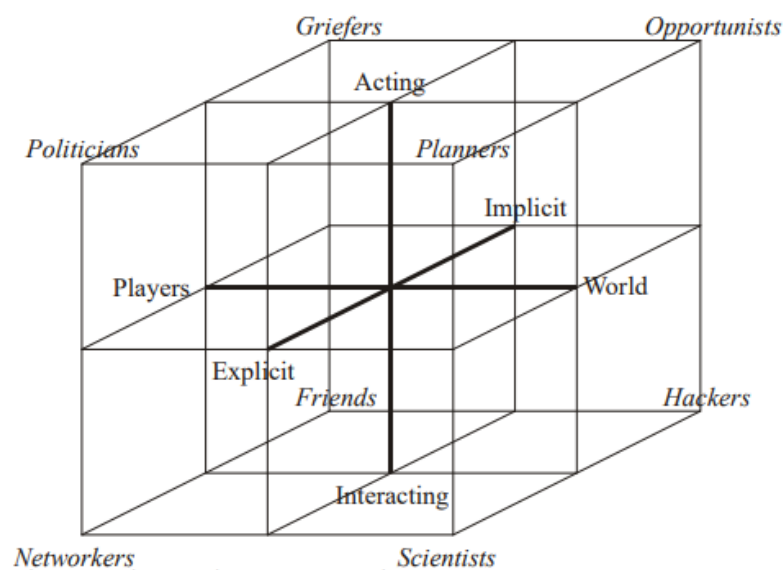


Figura 11 – Os oito novos tipos de jogadores de Bartle
Fonte: Bartle (2005)

Cada um dos 4 tipos originais foi subdividido em dois outros, cada um com sua forma implícita e explícita. Os *Killers* foram representados pelos *Griefers* e *Politicians*, *Opportunists* e *Planners* representam os *Achievers*, os *Explorers* são formados pelos *Hackers* e *Scientists*, enquanto *Networkers* e *Friends* formam os *Socializers* (BARTLE, 2003) (BARTLE, 2005). As características de cada um podem ser vistas abaixo:

Griefers são *Killers* implícitos, tendem a fazer qualquer coisa e utilizar de todos os meios possíveis para chamar atenção, ou atrapalhar outros jogadores/usuários (BARTLE, 2005).

Politicians são jogadores que buscam ter uma boa reputação com outros jogadores, porém não deixam de utilizar de manipulação ou realizar ações premeditadas. São a contraparte explícita dos *Killers* (BARTLE, 2005) (BARTLE, 2003).

Para os *Socializers* têm-se os *Networkers*, que são sua forma explícita e são os jogadores que tendem a interagir de forma aberta com quaisquer tipos de pessoas, sejam elas suas conhecidas ou não (BARTLE, 2003). Enquanto os *Friends* são os implícitos, e tendem a interagir primeiramente com jogadores conhecidos e que já possuem alguma conexão prévia. (BARTLE, 2003) (BARTLE, 2005).

Os *Scientists* e *Hackers* são os *Explorers* explícitos e implícitos respectivamente. O primeiro é o tipo de jogador que cria teorias, ou coloca em prática as teorias formadas por outros jogadores e explica novos fenômenos do jogo. Já o segundo refere-se aos jogadores que exploram o jogo por pura intuição, sem nenhum processo metódico de exploração, e são jogadores que buscam descobrir novos fenômenos do jogo (BARTLE, 2005).

Por fim, temos os *Opportunists* e *Planners*. O primeiro representa os *Achievers* implícitos, que tendem a aproveitar as oportunidades assim que surgem, mas também podem desistir de alguma coisa se algum obstáculo aparecer; estão sempre mudando de ideias rapidamente. O segundo são os *Achievers* explícitos, que impõem metas e as buscam; cada ação é pensada e realizada de acordo com um objetivo maior, tentam contornar quaisquer obstáculos que apareçam e não perdem o foco (BARTLE, 2005).

Além de ter novos tipos de jogadores, Bartle (2003) mostra que os tipos de jogadores podem transitar entre eles e, geralmente, seguem certo padrão. O padrão seguido geralmente se dá quando se alterna entre os tipos implícitos e explícitos de jogadores. Os jogadores começam com um dos tipos de comportamento, e conforme vão obtendo as informações necessárias ou alcançando objetivos maiores, começam a transitar para os outros tipos de jogadores. Um exemplo dessa transição é quando um jogador começa como um *Griefer* e, após algumas horas de jogo, quando está um pouco mais familiarizado com a jogabilidade e os objetivos, pergunta

para alguma outra pessoa ou procura saber de algum outro jeito, como passar de uma determinada parte, o que o deixa no momento como um *Networker*. Após obter essas informações e receber as recompensas, pode se tornar conhecido de uma pequena comunidade do jogo, e pode passar a influenciar essas pessoas, o deixando na condição de *Politician*. Por fim, o jogador já conhece bem sua base influenciável e acaba desenvolvendo uma amizade com alguns desses outros jogadores, o que o transforma em um *Friends* (BARTLE, 2005).

Um ponto de atenção nessa transformação do jogador é que isso acontece apenas para aquele jogo determinado que está jogando. Um jogador nunca será apenas de um desses tipos em quaisquer outros mundos virtuais.

3.2 AS MOTIVAÇÕES DOS JOGADORES

Yee (2006) é um dos pesquisadores que confrontam a teoria de Bartle e seus arquétipos, afirmando que não há uma tentativa sistemática para criar um *framework* da motivação de jogadores de MMORPG, e sim apenas um *framework* exploratório desenvolvido por Bartle, que servia apenas para MUD. O autor ainda traz a crítica a respeito de que a validação dos quatro tipos de jogadores de Bartle, (*Achiever, Explorer, Socializer, Killer*) e suas motivações não foram testados de forma empírica.

Para Yee, é importante saber a motivação dos jogadores de MMORPG devido às interações entre os usuários, o quanto se dedicam emocionalmente, como e se podem desenvolver habilidades interpessoais nesse tipo de jogo. Para conseguir essas respostas, foi desenvolvido uma série de questionários online para definir os dados demográficos, as motivações, e as suas experiências dentro do jogo. Durante um período de quatro anos (2000-2003), mais de 30.000 jogadores de MMORPG responderam os questionários, onde cada série continha entre 2.000 e 4.000 respostas de jogadores dos MMORPG mais jogados na época do estudo: *EverQuest*⁸, *Dark Age of Camelot*⁹, *Ultima Online*¹⁰, e *Star Wars Galaxies*¹¹ (YEE, 2006).

Um ponto interessante na pesquisa de Yee (2006) é a respeito dos resultados demográficos desses jogadores.

⁸ <https://www.everquest.com>

⁹ <http://darkageofcamelot.com/>

¹⁰ <https://uo.com/>

¹¹ Descontinuado em 15 de dezembro de 2011

A figura 12 mostra os resultados de uma população de 5.547 respostas:

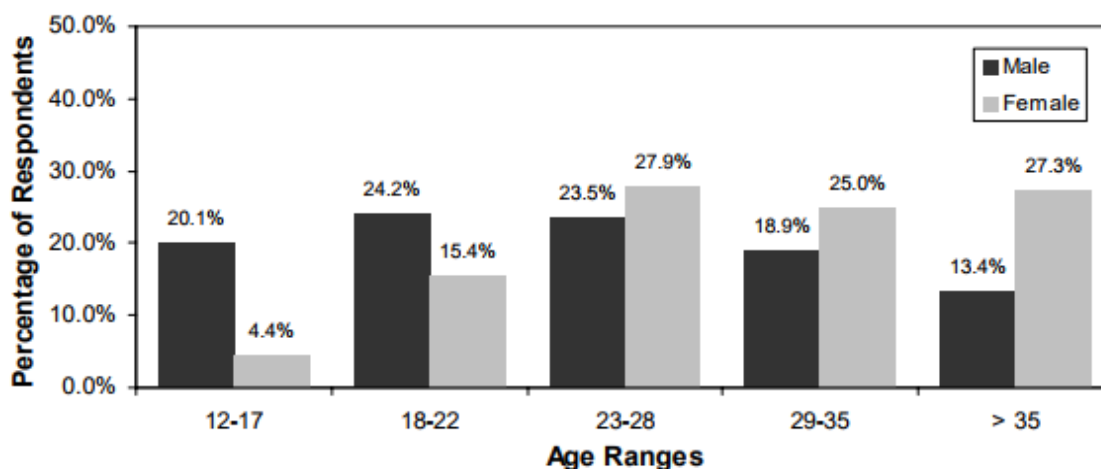


Figura 12 – Distribuição de Idade por gênero de jogadores (Homens = 4705, Mulheres = 788)
Fonte: Yee (2006)

A média de idade dos jogadores era de 26,57 anos. Enquanto a média do público feminino era de 31,72 anos, a do público masculino era 25,71, um pouco mais baixa que a média total. (YEE, 2006). Em questão à quantidade de horas jogadas pelos jogadores, foi encontrada uma média de 22,71 horas (total da população da pesquisa = 5471), como mostra a tabela 1 abaixo (YEE,2006).

Tabela 1 – Distribuição de horas jogadas por semana por gênero

Gender	0-10 hrs	10-20 hrs	20-30 hrs	30-40 hrs	40-50 hrs	50-60 hrs	> 60 hrs
Male (n = 4739)	24.5%	34.0%	20.6%	11.7%	4.6%	2.0%	1.8%
Female (n = 791)	22.6%	30.2%	23.3%	13.9%	5.5%	2.7%	1.9%

Fonte: Yee (2006)

Por mais pontos interessantes que tenham as diferenças entre os gêneros, não é o escopo dessa pesquisa, por isso as comparações entre gêneros não serão abordadas.

3.2.1 Motivações dos jogadores de Yee

As perguntas nos questionários de Yee tiveram como base os Arquétipos de Bartle (1996) e, a partir das respostas quantitativas do questionário, foram identificadas inicialmente 5 características que seriam as motivações de cada jogador: *Achievement*, *Relationship*, *Manipulation*, *Immersion* e *Escapism*. Os três primeiros podem ser relacionados aos de Bartle:

Achievement - *Achievers*, *Relationship* – *Socializer*, *Manipulation* – *Killers*, enquanto *Immersion* e *Escapism* não haviam sido identificados por Bartle (CAPLAN, WILLIAMS, YEE, 2008).

No primeiro momento não foram encontrados tipos de jogadores da categoria *Explorers* ou de qualquer tipo de exploração geográfica, por isso Yee utiliza das respostas qualitativas do questionário, e identifica mais 5 características, totalizando em 10 os componentes de motivação de jogadores (CAPLAN, WILLIAMS, YEE, 2008).

Esses 10 componentes foram divididos em 3 grandes grupos de motivação para jogadores: *Achievement*, *Social*, *Immersion*. A tabela 2 mostra como estão dispostas essas características por grupo (YEE, 2006).

Tabela 2 – Características de jogadores e seus agrupamentos

<i>Achievement</i>	<i>Social</i>	<i>Immersion</i>
Advancement	Socializing	Discovery
Progresso, Poder, Acumular, Status	Conversas Casuais, Ajudar Outros Fazer Amigos	Exploração, Conhecimento Encontrar Itens/Locais Escondidos
Mechanics	Relationship	Role-Playing
Números, Otimização Templates, Análises	Pessoal, Auto-Revelação, Encontrar e Dar Ajudas	Roteiro, História do Personagem Papéis, Fantasia
Competition	Teamwork	Customization
Desafiar outros jogadores, Provocação, Dominação	Colaboração, Grupos Recompensas em Grupo	Aparências, Acessórios, Estilo, Esquemas de Cores
		Escapism
		Relaxamento, Escapar da Vida Real Evitar Problemas da Vida Real

Fonte: Yee (2006)
Tradução: do autor

O grupo de *Achievement* relaciona as características de progresso no jogo, que contém os seguintes tipos:

- a) *Advancement*: São os jogadores que desejam avançar rapidamente no jogo, ganhar poderes e acumular itens ou status (YEE, 2006).
- b) *Mechanics*: Aqueles que têm interesse em entender cada linha das regras do mundo virtual para poder melhorar seu desempenho no jogo (YEE, 2006).
- c) Grupo que quer competir contra outros jogadores (YEE, 2006).

O segundo grupo é o *Social* que é dividido em:

- a) *Socializing*: Jogadores que gostam de ajudar, conversar, interagir com outros jogadores (YEE, 2006).
- b) *Relationship*: Quando se tem interesse de formar uma relação mais duradoura com outros indivíduos (YEE, 2006).
- c) *Teamwork*: Indivíduos que se satisfazem por participar de um grupo no mundo virtual e se sentir parte dele (YEE, 2006).

Por fim o grupo *Immersion* é composto de:

- a) *Discovery*: Querem descobrir e saber coisas que nenhum outro jogador sabe (YEE, 2006).
- b) *Role-Playing*: Jogadores que criam uma persona e começam a interpretar no jogo como se fossem a persona criada (YEE, 2006.)
- c) *Customization*: Tem interesse em customizar a aparência dos seus avatares (YEE, 2006).
- d) *Escapism*: Jogadores que usam do mundo virtual como forma de fugir de problemas do mundo real (YEE, 2006).

Com esse modelo com 10 características é possível cobrir o modelo de Bartle e, ainda, entender que o tipo Explorer de Bartle é composto de duas motivações não correlacionadas: *Discovery* nas explorações geográficas e *Mechanisms*, com a procura de como o sistema funciona e suas possíveis aberturas. Já com o segundo nível, com as 3 principais características, é possível classificar não só jogadores de MMORPG, como jogadores de outros tipos de jogos (CAPLAN, WILLIAMS, YEE, 2008).

Segundo Yee, Duchenaud e Nelson (2012), a taxonomia de Yee (2006) possui 3 fragilidades: ter um questionário muito longo para encontrar os 10 tipos de jogadores; a pesquisa foi realizada apenas no Ocidente, e por isso, deveria ter uma validação no Oriente, para verificar se é aplicável para todos os jogadores de MMORPG de outras culturas; e como o questionário correlaciona as respostas com o comportamento real no jogo.

Para resolver tais fragilidades, os autores realizaram aprimoramentos no modelo e testaram com jogadores de World of Warcraft (WoW). Para a primeira fraqueza, Yee, Duchenaud e Nelson, revisaram todas as perguntas do questionário original, e após ajustes semânticos nas perguntas e remoção de questões similares, chegaram a um total de 12 perguntas apenas, para representar 4 perguntas por grupo (Achievement, Social, Immersion), mostrando-se tão efetivo quanto o modelo anterior (YEE, DUCHENAUD, NELSON, 2012).

No caso de validar o modelo com jogadores não-ocidentais, o questionário foi traduzido para o chinês e enviado para jogadores de Hong Kong e Taiwan, pois possuem o mesmo servidor do jogo e também falam a mesma língua. O resultado foi que jogadores de uma cultura não-ocidental também são classificados nos 3 grandes grupos de jogadores de Yee (YEE, DUCHENAUT, NELSON, 2012).

Por fim, para sanar a terceira fragilidade, foi examinada a correlação entre o questionário e o comportamento dos jogadores selecionados na base de dados Armory do WoW, onde é possível encontrar informações de jogabilidade de um determinado personagem, desde o dia que foi criado, em algumas categorias pré-definidas (YEE, DUCHENAUT, NELSON, 2012), (YEE et al., 2011). Os jogadores foram analisados pelas seguintes categorias: *Quests*, *Exploration*, *PvP (Player-vs-player)*, *Dungeons/Raids*, *Professions* e *World Events*, pois são atividades que eram calculadas automaticamente e não demandavam desenvolvimento para obter as informações (YEE, DUCHENAUT, NELSON, 2012). O resultado obtido foi de que os itens *Quests*, *Exploration* e *Professions* tendem a jogadores que estão no grupo *Immersion*, devido às características de *Discovery* e *Role-Playing*. *PvP* e *Dungeons* correlacionam-se com o grupo *Achievement*, devido às características *Advancement* e *Competition*. *Dungeons* ainda se correlaciona com o *Social*, por ser uma atividade que se faz em grupos; logo, a característica *Teamwork* está presente nesse item. Por fim, *World Events* correlaciona-se com os do grupo *Immersion*, devido ao fato de serem eventos sazonais e mais voltados ao mundo virtual de WoW, privilegiando os jogadores que tendem mais à característica de *Role-Playing* (YEE, DUCHENAUT, NELSON, 2012).

Os tipos de jogadores de Yee podem ser aplicados para jogadores de outros tipos de MMORPG, como é visto em Yee (2006) e Yee, Duchenaute, Nelson (2012). Porém, ele é sempre validado através de um questionário que os próprios usuários devem responder previamente, e podem ser encontrados com uma análise do *gameplay* dos jogadores no mundo virtual escolhido.

3.3 USO DE DADOS DE TELEMETRIA DE JOGOS PARA DEFINIR TIPOS DE JOGADORES

Com o uso de telemetria de dados, as classificações dos jogadores através de questionários são descartadas, e apenas com os dados obtidos através da telemetria é que se faz a análise e o agrupamento dos tipos de jogadores. Nessa linha de pesquisa tem-se dois trabalhos que se destacam, os de Drachen, Canossa, Yannakakis (2009) e Drachen et al. (2012).

Ambos estudos utilizam de redes neurais não supervisionadas, redes em que não há necessidade de conhecer a saída desejada, aprendendo apenas com os dados de entrada (MACIEL, 2008), (KOHONEN, 2001).

A diferença é que em Drachen, Canossa, Yannakakis (2009) utiliza-se uma rede *Self-Organized Maps* (SOM) e em Drachen et al. (2012) utiliza-se dos algoritmos K-means e *Simplex Volume Maximization* (SIVM).

Esses tipos de jogadores só puderam ser identificados através da colaboração dos estúdios desenvolvedores em fornecer os dados de telemetria de jogos. Por isso, esse modelo é muito difícil de ser realizado, uma vez que não há sempre colaboração dos estúdios com a academia (DRACHEN, CANOSSA, YANNAKAKIS, 2009) (DRACHEN ET AL., 2012).

3.3.1 Tipos de jogadores em jogos de plataforma

O estudo de Drachen, Canossa, Yannakakis (2009), mostra-se relevante por ser um dos primeiros com dados de telemetria, fornecidos por uma desenvolvedora de jogos. Os dados dos jogadores foram coletados do jogo *Tomb Raider: Underworld*, durante o ano de 2008. As informações foram obtidas através de uma ferramenta de coleta da própria desenvolvedora (EIDOS), e eram coletadas diretamente de jogadores, ao invés de em laboratório.

Após a coleta, os dados foram agrupados através de uma rede neural *Self-Organized Maps* (SOM). A arquitetura dessa rede neural é disposta em uma grade de dimensão baixa (2 ou 3) e é definida com um peso sináptico $\vec{m} = \{w_1, \dots, w_m\}$, onde m é o tamanho do mapa neural. Cada neurônio é conectado com seu vizinho de forma lateral ou diagonal, e suas formas mais utilizadas são a quadrada e a hexagonal (DRACHEN, CANOSSA, YANNAKAKIS, 2009) (MACIEL, 2008) (KOHONEN, 2001).

O jogo escolhido para o estudo de Drachen, Canossa e Yannakakis (2009) é do tipo plataforma em terceira pessoa, no qual os jogadores precisam resolver quebra-cabeças para avançar, além de ter que explorar o mapa em busca de itens colecionáveis, enfrentar inimigos dos mais variados tipos (animais, pessoas, monstros), tomando o devido cuidado para não cair de penhascos ou se afogar nos rios percorridos.

Foram coletados dados de 1365 jogadores, e as *features* escolhidas foram: Causas de morte (divididas em três causas: Oponentes, Ambiente, Quedas), Número total de mortes, Tempo de conclusão do jogo, e Ajudas necessárias (DRACHEN, CANOSSA, YANNAKAKIS, 2009).

Após a normalização dos valores dos dados coletados, foram encontrados quatro grupos pela rede neural SOM. O primeiro grupo foi chamado de *Veterans*: jogadores que morrem

poucas vezes, a maior parte dessas mortes decorrentes do Ambiente e completam o jogo mais rápido que os outros tipos (DRACHEN, CANOSSA, YANNAKAKIS, 2009).

Jogadores do *cluster 2* são denominados *Solvers*, aqueles que demoram mais tempo na conclusão do jogo, morrem mais vezes devido a Quedas, não utilizam tanto o recurso Ajuda, além de sua baixa taxa de morte por Oponentes, e pelo Ambiente demonstram que os jogadores se movimentam com cautela (DRACHEN, CANOSSA, YANNAKAKIS, 2009).

Pacifists são os jogadores do grupo 3. São os que mais morrem devido aos Oponentes, mas seu tempo de conclusão do jogo é abaixo da média, o que demonstra sua habilidade (DRACHEN, CANOSSA, YANNAKAKIS, 2009).

O último grupo é representado pelos *Runners*, jogadores que morrem com frequência, principalmente por Oponentes e pelo Ambiente, porém, possuem um baixo tempo de conclusão do jogo, e um alto índice de Ajuda para completar mais rapidamente os desafios deste (DRACHEN, CANOSSA, YANNAKAKIS, 2009).

3.3.2 Tipos de jogadores em MMORPG e Jogos *First Person Shooter* (FPS)

O segundo exemplo de uso de redes não supervisionadas para classificação de jogadores é o de Drache et al. (2012), onde foram coletados 260.000 dados de telemetria dos jogos *Battlefield 2: Bad Company 2*¹² (BF2sBC2) – um jogo *multiplayer* de guerra, – e do MMORPG *Tera*¹³, e utiliza-se de métodos não supervisionados *k-means* e *Simplex Volume Maximization* (SIVM) para agrupar os tipos de jogadores de cada jogo. (DRACHEN et al., 2012).

Para o jogo Tera foi escolhido um *dataset* de 250.000 dados e foram analisados dois grupos de funcionalidades: habilidades dos personagens e *gameplay*. Nas habilidades dos personagens estão suas características, como: raça, classe, ataque, defesa, entre outras. A análise de *gameplay* tem as características de como cada jogador joga. É aqui que se tem as informações de quantidade de monstros derrotados, amigos no jogo, missões concluídas. É essa a sessão que interessa ao presente trabalho. (DRACHEN et al., 2012).

Do *dataset* fornecido pela desenvolvedora do Tera, foram escolhidas as seguintes características de jogo: quantidade de missões completadas, número de amigos, quantidade de *achievements* do jogo, nível da habilidade de *Mining* e *Plants*, número de inimigos mortos, quantidade de itens coletados durante o jogo, número de vezes que o jogador foi morto por um

¹² <https://www.battlefield.com/pt-br/games/battlefield-bad-company-2>

¹³ <http://tera.enmasse.com/>

monstro, número de vezes que um jogador criou ou comprou algo de um leilão, nível do jogador (DRACHEN et al., 2012).

No caso do *Battlefield 2: Bad Company 2*, foram selecionados dados de 10.000 jogadores da base do servidor que coleta as informações de telemetria dos jogadores. Diferente do Tera, que é um jogo MMORPG, BF2BC2 é um jogo de tiro em primeira pessoa, com suporte a multiplayer de até 32 jogadores simultâneos. Diferente de um MMORPG, onde jogadores têm uma vasta gama de classes e raças de personagens, no BF2BC2 jogadores assumem personagens com classes limitadas a *Assault*, *Medic*, *Recoon*, *Engineer*, fixos, onde cada um tem uma especialidade. Por exemplo, *Recoons* possuem um kit de armas do tipo *sniper* com longo alcance (DRACHEN et al., 2012).

Foram selecionados dados de um grupo de 10.000 jogadores, de uma base que coleta telemetria dos jogadores, e um total de 11 características foram escolhidas para serem analisadas: Total de pontos, Pontos de experiência dos jogadores (*Skill level*), Total de horas jogadas, Percentual *Kill/Death*, Precisão dos tiros, Pontuação por minuto, Mortes por minuto ou *Kills* por minuto, Rodadas jogadas, Pontos ganhos e o número de *kills* e mortes por classe, Tempo total gasto nos veículos disponíveis (DRACHEN et al., 2012).

Após normalizar os dados de cada jogo, ambos foram analisados com os algoritmos *K-means* e *SIVM*. Esses algoritmos foram escolhidos pois, enquanto o primeiro pode ser utilizado para obter informações do comportamento em geral dos jogadores no ambiente virtual, o segundo é utilizado para obter informação sobre comportamentos extremos (DRACHEN et al., 2012).

3.3.2.1 Tipos de jogadores do Tera

Para o jogo Tera, os jogadores foram divididos em 4 grupos, baseados em seus níveis (1-10, 11-20, 21-30, 30-32). No trabalho, optou-se por mostrar os grupos de jogadores do último grupo, por serem os de maior nível no jogo e apresentarem o maior número de tipos de jogadores. Foram identificados 6 grupos de jogadores com o *k-means*: *Elite*, *Stragglers*, *Average Joes*, *The Dependables*, *Worker I* e *Worker II*. Com o *SIVM* foram identificados também 6 grupos: *Elite*, *Stragglers*, *Planters*, *Miners*, *Auction Devils* e *Friendly Pros* (DRACHEN et al., 2012).

Nos dois casos, os “Elite” são jogadores com maior pontuação em todas as categorias, exceto na categoria de mortes por monstros do jogo, e baixo nível nas *skills* de *Mining* e *Plants*, o que mostra que são focados em desempenho no jogo (DRACHEN et al., 2012).

Os *Stragglers* também foram identificados nos dois algoritmos. Os jogadores dessa categoria são os que possuem a menor pontuação em todas as categorias do jogo, inclusive mortes por monstros (DRACHEN et al., 2012).

Os *Average Joes* apresentam a maior pontuação entre os piores jogadores (DRACHEN et al., 2012).

Os próximos tipos foram identificados apenas pelo *k-means*. Os *Dependables* possuem uma pontuação mediana em todas as categorias, um número maior de amigos no jogo, e têm a segunda maior quantidade de monstros matados (DRACHEN et al., 2012).

O grupo *Worker I* e *Worker II* são correspondentes aos *Average Joes* e *Dependables*, porém *Worker I* possuem um foco nas habilidades de *Mining* e *Plants*. Os *Workers II* possuem o maior nível das habilidades de *Mining* e *Plants* do jogo, além de uma grande quantidade de itens coletados (DRACHEN et al., 2012).

O quadro 1, a seguir, mostra o percentual incidente de cada grupo na amostra coletada.

Quadro 1 – Grupos, percentual e as características dos jogadores identificados no Tera, por meio do algoritmo k-means.

Title	%P	Characteristics
Elite	5.78	Highest scores for all features except Mining and Plants which are the lowest in the game.
Stragglers	39.4	Lowest scores for all features, including deaths from monsters.
Average Joes	12.7	Better scores across all categories than Low Performers, 4 th ranked overall
Dependables	18.6	Average scores across all categories, high number of friends, 3 rd ranked overall, 2 nd rank in monster kills
Worker I	15.9	Similar to the Average Joes, but high Mining and Plants, and loot value 3 rd ranked.
Worker II	7.6	Similar to The Dependables, but highest Mining and Plants value in the game. 2 nd ranked overall. Loot 2 nd ranked.

Fonte: Drachen et al., 2012.

As características dos jogadores identificados apenas pelo SIVM são como demonstradas. Os *Planters* e *Miners* são os que tem maiores pontuações nas suas respectivas habilidades. *Auction Devils* são aqueles jogadores que utilizam mais as casas de leilão em busca de itens e possuem um certo nível na habilidade *Mining*, muito provavelmente para fabricar recursos para os leilões. Já os *Friendly Pros* possuem um comportamento similar aos do grupo anterior, porém não possuem uma assiduidade tão grande em casas de leilão, como pode ser visto no quadro 2 (DRACHEN et al. 2012).

Quadro 2 – Grupos, percentual e as características dos jogadores identificados no Tera através do algoritmo SVM.

Title	%P	Characteristics
Elite	3.9	High scores overall, except for Mining/Plants and deaths from monsters. No auctions created.
Stragglers	7.6	Low scores overall, dies a lot from monsters
Planters	21.6	Middling scores, but high Plants skill
Miners	15.0	Middling scores, but high Mining skill
Auction Devils	1.1	Highest auction and achievement score. 2 nd ranked loot and kills scores. 2 nd ranked friends score, high mining score
Friendly Pros	50.8	Highest friends score, scores similar to Auction Devils apart from low auction score and 2 nd lowest loot score

Fonte: Drachen et al., 2012.

3.3.2.2 Tipos de jogadores do *Battlefield 2: Bad Company 2*

Após analisar 22 tipos de dados, os jogadores do BF2BC2 foram divididos em 7 categorias diferentes, e, similar aos tipos de jogadores do Tera, os dois algoritmos foram consistentes em seus agrupamentos. Com o *k-means* foram identificados: *Assassins*, *Veterans*, *Target Dummies*, *Snipers*, *Soldiers*, *Assault-Engineer* e *Trainee veterans*. Com o SVM foram identificados: *Assassins*, *Veterans*, *Target Dummies*, *Assault-Recon*, *Medic-Engineer*, *Assault "specialist"* e *Driver Engineers* (DRACHEN et al., 2012).

É notado que novamente têm-se categorias que se repetem (*Assassins*, *Veterans*, *Target Dummies*), enquanto as outras categorias são quase todas voltadas às classes e suas especialidades, em exceção os *Trainee veterans*.

Os *Assassins* são jogadores que têm a maior taxa Kill/Death entre todas as classes disponíveis, a maior taxa de *Kills* por minuto, porém não é a primeira em quantidade de horas jogadas (DRACHEN et al., 2012).

Veterans são os jogadores que estão em segundo lugar em quase todas as categorias dos *Assassins*, porém possuem a maior quantidade de horas jogadas (DRACHEN et al., 2012).

Os opostos desses dois jogadores são os *Target Dummies*. Esses jogadores possuem uma baixa pontuação em todas as categorias, baixa pontuação por minuto, não têm a tendência de jogar muito, e são alvos fáceis para qualquer tipo de jogador um pouco mais experiente (DRACHEN et al., 2012).

Os outros tipos de jogadores estão mais correlacionados às classes de jogo, portanto, quanto mais se joga com uma classe, maior a chance de ser classificado nesses tipos de jogadores.

A exceção feita nessa classificação são os *Trainee veterans*, que apresentam números abaixo dos *Veterans*, porém com menor tempo de jogo (DRACHEN et al., 2012).

Os quadros 3 e 4, abaixo, mostram as características e os percentuais de jogadores encontrados com cada um dos algoritmos.

Quadro 3 – Grupos, percentual e características dos jogadores identificados no BF2BC2, através do algoritmo k-means.

Title	%P	Characteristics
Snipers	7.4	Median SpM, overall low-middling values, high DpM, extremely high Accuracy. Highest kit score is Medic.
Soldiers	27.9	Median SpM, overall low-middling values, high DpM, highest kit score is Assault.
Assault-Engineer	13.1	Similar to Soldiers but better skill value, high Engineer and Assault scores and K/D ratios.
Target Dummies	26.0	Lowest scores for all values (including Playtime) except high DpM.
Trainee veterans	10.7	Comparable to Veterans, but 2 nd rank in most features, and lower Playtime.
Assassins	10.9	Highest rank in all K/D-ratios, highest KpM, low Playtime, low DpM.
Veterans	4.1	High Playtime, 2 nd rank in most features, highest overall Skill level.

Fonte: Drachen et al., 2012.

Quadro 4 – Grupos, percentual e características dos jogadores identificados no BF2BC2, através do algoritmo SIVM.

Title	%P	Characteristics
Assault-Recon	1.4	High KpM and DpM, low Accuracy, average SpM, 2nd highest K/D overall.
Medic-Engineer	0.8	High Vehicle time, Skill level and Accuracy, 2nd highest SpM.
Assault "specialist"	5.0	Focus on Assault, but low score, high DpM and Playtime, low Skill, K/D and Accuracy.
Driver Engineers	1.1	Extremely high Vehicle time, high Playtime, Score and Accuracy, 2nd highest K/D, lowest DpM, low KpM.
Assassins	61.6	Highest K/D, high KpM, lowest Playtime, very low DpM.
Veterans	2.01	Highest Score, Playtime and Rounds played, overall high values
Target Dummies	28.1	Extremely low K/D, lowest Skill, SpM and KpM min. scores for all features but Playtime and Rounds played.

Fonte: Drachen et al., 2012.

3.4 OS TIPOS DE JOGADORES PARA AMBIENTES GAMIFICADOS DE MARCZEWSKI

Os tipos de jogadores de Marczewski (2015) foram baseados nos 8 tipos de jogadores de Bartle (2003). A diferença é que Marczewski utiliza a palavra usuário e sistema, ao contrário

de jogadores e mundo de jogo, pois seus tipos de jogadores são para ambientes gamificados e não para jogos. A figura 13 mostra os oito tipos de jogadores.

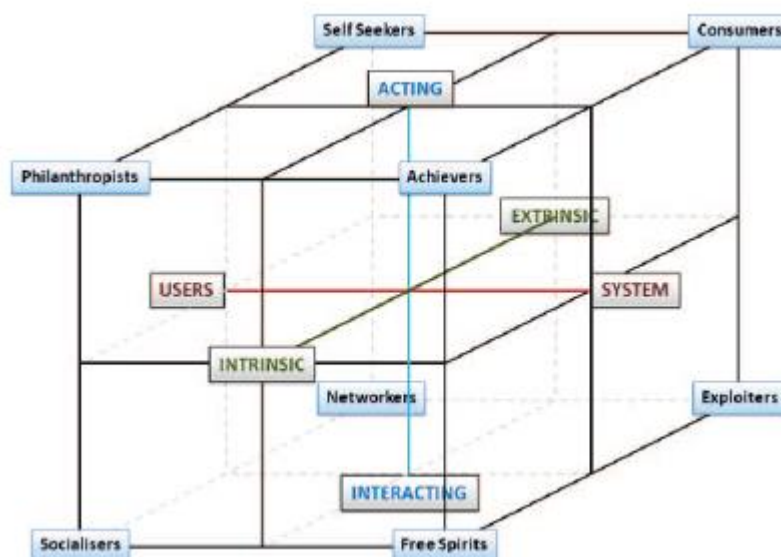


Figura 13 – 8 tipos de jogadores de Marczewski
Fonte: Hebert et al., 2014

A Figura 13 exibe o primeiro esboço de Marczewski sobre os tipos de jogadores e, com o auxílio do modelo motivacional RAMP, o autor aperfeiçoou seu modelo e chegou a um com 6 tipos de jogadores (MARCZEWSKI, 2015).

RAMP é a base motivacional para os tipos de usuários de Marczewski, é a sigla em inglês para *Relatedness, Autonomy, Mastery, Purpose* (MARCZEWSKI, 2013).

Relatedness é tudo que envolve os relacionamentos, como se dá a conexão dos usuários no ambiente virtual. Para Marczewski (2013), esse é o principal motivador de um ambiente gamificado. Se os usuários se sentem parte de alguma coisa, eles usam com muito mais afinco.

Muitas pessoas não gostam de se sentir controladas, ou que são limitadas. Criando ambientes pensando no *Autonomy*, os usuários poderão ter a sensação de liberdade e autonomia para utilizar o ambiente do jeito que lhe for mais prazeroso (MARCZEWSKI, 2013).

Mastery é o processo de atingir o nível máximo em algo. É um processo gradual que demanda certo tempo e, para manter os usuários motivados a atingir o domínio, é importante que os desafios aumentem conforme o conhecimento sobre o ambiente também aumenta (MARCZEWSKI, 2013).

A última motivação é *Purpose*, que é o que leva um usuário a continuar usando um sistema, um jogador a continuar frequentando um determinado mundo virtual. Podem ser

identificados aqui os usuários/jogadores que estão no ambiente apenas para ajudar outros por ajudar (MARCZEWSKI, 2013).

Utilizando o RAMP como base, Marczewski chegou a 6 tipos de jogadores: *Socializers*, *Free Spirits*, *Achievers*, *Philanthropists*, *Players* e *Disruptors* (MARCZEWSKI, 2015).

A personalidade dos *Socializers* é similar aos de Bartle (1996), são jogadores que querem interagir com outros jogadores, e se interessam pela parte do sistema que as possibilita realizar esse tipo de interação. São motivados pelo *Relatedness* (MARCZEWSKI, 2015).

Os *Free Spirits* são tanto aqueles que querem explorar o ambiente sem restrições e encontrar a maior quantidade de aberturas possíveis, quanto criar novas funcionalidades ou customizar o que é possível do ambiente em que se encontra. Estão ligados ao *Autonomy*, e sempre buscam uma forma criativa de expressão (MARCZEWSKI, 2015).

A principal motivação dos *Achievers* é ao menos obter coisas do ambiente e, se possível, ser o melhor em todas as opções disponíveis. São usuários que irão competir com os outros, mas com a intenção de serem os melhores, e não se gabar disso. Com isso, um sistema que irá agradar os que possuem esse perfil é um que potencialize o *Mastery* (MARCZEWSKI, 2015).

Os que se encaixam como *Philanthropists* querem se sentir parte de algo, querem ajudar outras pessoas sem esperar nada em troca. O sistema que agrada mais esse perfil é aquele que tem como foco no *Purpose* (MARCZEWSKI, 2015).

Players e *Disruptors* são os tipos que não estão ligados ao RAMP, mas que estão presentes em qualquer ambiente, independentemente do sistema. *Players* são os primeiros a aderirem ao sistema em questão, estão atrás de recompensas, de reconhecimento, só que, uma vez que as recompensas acabam, são os primeiros a sair do ambiente e procurar outro que atenda esses requisitos (GAMIFICATION WORLD CONGRESS, 2014) (MARCZEWSKI, 2015).

Disruptors buscam mudança, irão utilizar de todas as formas possíveis para alterar o sistema, ou agir e encontrar como se aproveitar de possíveis brechas no mesmo. É o grupo que Marczewski cita no Gamification World Congress (2014), ou os elimine do sistema, ou os dê algo para fazer. Se os participantes deste grupo sentirem que têm poder para realizar mudanças, serão de grande ajuda para o sistema. São os usuários que, se sentirem que podem ajudar a modificar, que contribuirão para o ambiente (GAMIFICATION WORLD CONGRESS, 2014) (MARCZEWSKI, 2015). A Figura 14, abaixo, exhibe a categorização de Marczewski, com os detalhes de cada arquétipo sistematizados.

Os tipos de jogadores são constantemente validados pelo autor, através de um questionário online¹⁴. O autor revisa os tipos de usuários de tempos em tempos, e valida sua teoria conforme as respostas do questionário (MARCZEWSKI, 2013).

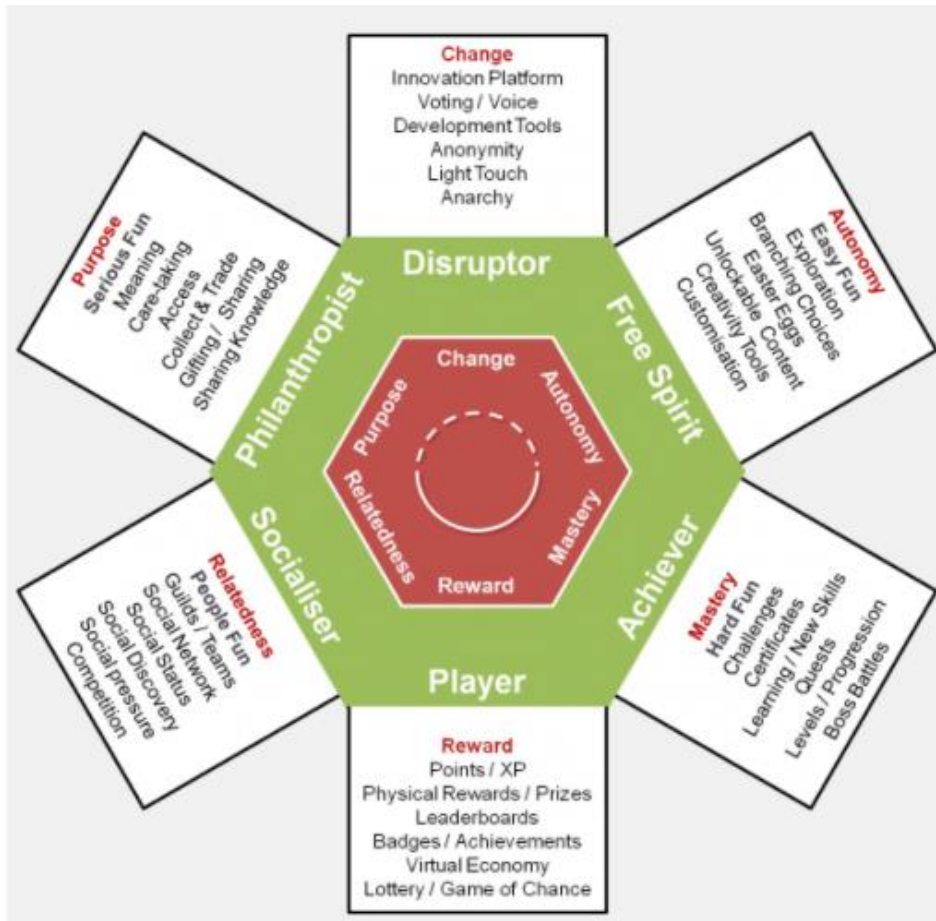


Figura 14 – Tipos de usuários de Marczewski
Fonte: Marczewski, 2015

3.5 SOBRE TIPOS DE JOGADORES

Conhecer os tipos de jogadores tornou-se fundamental para obter sucesso nos ambientes virtuais analisados. Quanto maior o conhecimento, melhor será o desenvolvimento, e os tipos desejados terão o foco necessário.

É importante notar que cada autor desenvolve um tipo de taxonomia diferente. Enquanto Yee baseia-se primordialmente em resultados de pesquisa entre jogadores, Marczewski utiliza seus questionários para validar sua pesquisa. Bartle não utiliza questionários, porém sua

¹⁴ Disponível em: <https://gamified.uk/UserTypeTest2016/user-type-test.php>

pesquisa fora baseada em discussões diretas com jogadores do ambiente que criou. Enquanto Drachen et al. (2012) utiliza dados reais de jogadores no seu ambiente natural: os jogos. Esse último aparenta obter a melhor classificação dos jogadores, porém com a dificuldade de se obter dados reais, nem sempre é possível realizar esse tipo de classificação.

É importante ressaltar que a intenção deste trabalho não é definir qual o melhor tipo de classificação, e sim mostrar quais são os tipos de jogadores mais frequentes na época da pesquisa.

4 RESULTADOS E DISCUSSÃO

Para realizar esse experimento, foi escolhido o jogo *World of Warcraft* (WoW), pois ele se encaixa nas definições de um jogo MUD: *Jogo multiplayer online*, ambientado em um mundo virtual, jogado em tempo real (THE STATISTICS PORTAL, 2015). Apesar de não ter mais a popularidade de seus anos iniciais, WoW ainda é um jogo com muitos jogadores ativos: 5,5 milhões de jogadores no final do terceiro trimestre de 2015, segundo a última publicação oficial da Blizzard (produtora e distribuidora oficial do jogo) (BARROS; NOTARGIACOMO, 2016), como pode ser visto na Figura 15. Jogos MMORPG têm perdido espaço para os jogos MOBA (*Multiplayer Online Battle Arena*), porém continuam tendo popularidade no seu nicho.

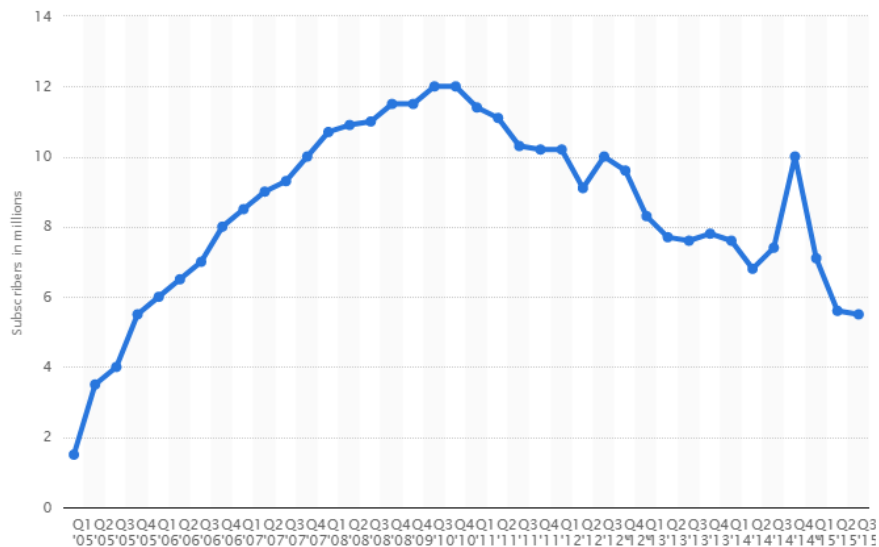


Figura 15 – Número de jogadores ativos no jogo World of Warcraft de janeiro 2005 a setembro 2015, divulgado pela Blizzard

Fonte: Barros, Notargiacomo (2016)

O grupo de dados escolhido foi o World of Warcraft Avatar History (WoWAH), que foi extraído do jogo durante um período de 3 anos (janeiro de 2006 a janeiro de 2009), fornecendo dados de um total de 91.605 avatares (LEE et al. 2011).

Este trabalho pretende mostrar como as informações contidas nessa extração de dados podem ajudar a combinar e agrupar os jogadores nos arquétipos de Bartle e, com isso, definir quais são as melhores formas de encontrar esses jogadores através de seu game play, com ajuda de uma rede neural.

4.1 O GRUPO DE DADOS

Os seguintes dados do WoWAH foram coletados e armazenados durante os três anos do período de observação: data e horário da captura dos dados, sequência da coleta de dados, ID do Avatar, *Guild*, Nível, Raça, Classe e Zona do jogo em que o jogador se encontra. Esses dados foram coletados do servidor *Light's Hope* de *Taiwan*. O jogo ainda possui duas facções: *Alliance* e *Horde*, os dados foram extraídos da facção *Horde* (LEE et al., 2011).

Para esse estudo, foi selecionado um grupo de dados reduzido, por ser um primeiro ensaio. Os dados são do ano de 2008, mostrando as informações de todos os seus usuários naquele ano (KAGGLE, 2015).

4.2 PREPARAÇÃO DOS DADOS

Para iniciar o ensaio, foi necessário realizar algumas alterações nos dados fornecidos e criar algumas colunas adicionais, como mostrado nos quadros 5 e 6 e nas figuras 16 e 17. A figura 16 mostra como a tabela de dados é representada na sua forma original (como descrita no quadro 5) e a 17, mostra como ficam os dados após os ajustes necessários (descritos no quadro 6).

Quadro 5 – Dados iniciais disponíveis na extração reduzida do WoWAH

Campo	Valores
<i>Char</i>	Inteiro > 0
<i>Level</i>	Inteiro > 0
<i>Race</i>	Raças disponíveis no jogo
<i>Charclass</i>	Classes disponíveis no jogo
<i>Zone</i>	Uma das 229 zonas do WoW
<i>Guild</i>	Inteiro > 0
<i>Timestamp</i>	Jan, 2008 a Jan, 2009

Fonte: LEE et al., 2011.

Quadro 6 – Dados do WoWAH ajustados para o experimento

Campo	Valores
<i>Char</i>	Inteiro > 0
<i>Level</i>	Inteiro > 0
<i>Race</i>	Raças disponíveis no jogo

<i>Charclass</i>	Classes disponíveis no jogo
<i>Zone</i>	Uma das 229 zonas do WoW
<i>Guild</i>	Inteiro > 0
<i>Timestamp</i>	Jan, 2008 a Jan, 2009 com horas
<i>Current Date</i>	Data separada das horas
<i>Hour</i>	Inteiro entre 0 e 23
<i>Month</i>	Jan a Dez
<i>Month Index</i>	Inteiro de 1 a 12
<i>Activation Date</i>	Data de ativação do jogador – de Jan 2008 a Jan 2009
<i>DSI</i>	<i>Days Since Install</i> (Dias desde a instalação) inteiro 0 ou 1
<i>Type</i>	Agrupamento das Raças e Classes que um mesmo jogador pode ter

Fonte: LEE et al., 2011

	char	level	race	charclass	zone	guild	timestamp	date
1	59425	1	Orc	Rogue	Orgrimmar	165	01/01/08 00:02:04	2008-01-01
2	65494	9	Orc	Hunter	Durotar	-1	01/01/08 00:02:04	2008-01-01
3	65325	14	Orc	Warrior	Ghostlands	-1	01/01/08 00:02:04	2008-01-01
4	65490	18	Orc	Hunter	Ghostlands	-1	01/01/08 00:02:04	2008-01-01
5	2288	60	Orc	Hunter	Hellfire Peninsula	-1	01/01/08 00:02:09	2008-01-01
6	2289	60	Orc	Hunter	Hellfire Peninsula	-1	01/01/08 00:02:09	2008-01-01
7	61239	68	Orc	Hunter	Blade's Edge Mountains	243	01/01/08 00:02:14	2008-01-01
8	59772	69	Orc	Warrior	Shadowmoon Valley	35	01/01/08 00:02:14	2008-01-01
9	22937	69	Orc	Rogue	Warsong Gulch	243	01/01/08 00:02:14	2008-01-01
10	23062	69	Orc	Shaman	Shattrath City	103	01/01/08 00:02:14	2008-01-01
11	48432	70	Orc	Warrior	Blade's Edge Mountains	79	01/01/08 00:02:19	2008-01-01
12	582	70	Orc	Warrior	Sethikk Halls	19	01/01/08 00:02:19	2008-01-01

Showing 1 to 12 of 10,826,734 entries

Figura 16 – Dados iniciais do WoWAH

Fonte: Criação do autor

	char	level	race	charclass	zone	guild	timestamp	current.date	hour	month	month.idx	activation.date	dsi	type
1	59425	1	Orc	Rogue	Orgrimmar	165	2008-01-01 00:02:04	2008-01-01	0	jan	1	2008-01-01	0	59425.Orc.Rogue
2	65494	9	Orc	Hunter	Durotar	-1	2008-01-01 00:02:04	2008-01-01	0	jan	1	2008-01-01	0	65494.Orc.Hunter
3	65325	14	Orc	Warrior	Ghostlands	-1	2008-01-01 00:02:04	2008-01-01	0	jan	1	2008-01-01	0	65325.Orc.Warrior
4	65490	18	Orc	Hunter	Ghostlands	-1	2008-01-01 00:02:04	2008-01-01	0	jan	1	2008-01-01	0	65490.Orc.Hunter
5	2288	60	Orc	Hunter	Hellfire Peninsula	-1	2008-01-01 00:02:09	2008-01-01	0	jan	1	2008-01-01	0	2288.Orc.Hunter
6	2289	60	Orc	Hunter	Hellfire Peninsula	-1	2008-01-01 00:02:09	2008-01-01	0	jan	1	2008-01-01	0	2289.Orc.Hunter
7	61239	68	Orc	Hunter	Blade's Edge Mountains	243	2008-01-01 00:02:14	2008-01-01	0	jan	1	2008-01-01	0	61239.Orc.Hunter
8	59772	69	Orc	Warrior	Shadowmoon Valley	35	2008-01-01 00:02:14	2008-01-01	0	jan	1	2008-01-01	0	59772.Orc.Warrior
9	22937	69	Orc	Rogue	Warsong Gulch	243	2008-01-01 00:02:14	2008-01-01	0	jan	1	2008-01-01	0	22937.Orc.Rogue
10	23062	69	Orc	Shaman	Shattrath City	103	2008-01-01 00:02:14	2008-01-01	0	jan	1	2008-01-01	0	23062.Orc.Shaman
11	48432	70	Orc	Warrior	Blade's Edge Mountains	79	2008-01-01 00:02:19	2008-01-01	0	jan	1	2008-01-01	0	48432.Orc.Warrior
12	582	70	Orc	Warrior	Sethikk Halls	19	2008-01-01 00:02:19	2008-01-01	0	jan	1	2008-01-01	0	582.Orc.Warrior

Figura 17 – Dados adaptados para o experimento

Fonte: Criação do autor

O experimento foi dividido em duas partes: Inicialmente foi analisado um grupo controlado de jogadores e, para isso, foi escolhida a maior *guild* do jogo. Esse grupo serviu para realizar análises nos dados, e como eles poderiam ser trabalhados para que o objetivo final fosse obtido.

Após a realização do experimento inicial, foi utilizada toda a base de dados para verificar se os resultados do grupo reduzido se aplicavam ao conjunto de todos os jogadores.

Para realizar o experimento, foi necessário que o pesquisador realizasse observação prévia da forma como os jogadores se comportam no ambiente virtual. Foi criada uma conta no WoW com essa finalidade e, durante o período de 3 meses de observação, foi possível capturar alguns padrões de comportamento que serão comentados durante essa seção.

Essas observações eram realizadas diariamente entre os horários das 22:00 às 00:00 e de sexta-feira e sábado até às 2:00, por serem os dias em que os jogadores possuem mais tempo livre para jogar.

As observações foram realizadas em cinco regiões: Capitals, Battlegrounds, Outlands, Northrend e Arena. Apesar do jogo ter evoluído e possuir novas regiões, essas eram as regiões disponíveis à época da extração disponível. Foi possível observar que cada uma dessas regiões possui suas particularidades, o que atrai um tipo distinto de estilo de jogador.

As Capitals são as cidades onde ficam as bases de cada facção e onde os líderes (não jogáveis) ficam para liberar novas missões ou habilidades. Essas cidades possuem também bancos, caminhos diretos para outras regiões e cidades, caixa de mensagens, casa de leilões, treinadores de classe e raça, o que as torna uma região de interação social e econômica. Aqui, por ser uma região livre de combate entre jogadores ou contra inimigos, jogadores passam a maior parte do tempo interagindo entre si ou com os *non-playable characters* (personagens não jogáveis: NPC), resolvendo suas melhorias de raça ou classe de personagem, lendo as mensagens para coletar os itens de alguma missão completada, comprando novos itens de melhoria ou repondo seu estoque de poções.

A segunda região observada são os Battlegrounds. Essa é uma região dedicada ao PvP de grupos e só se pode entrar quando se atinge o nível 10. Há duas formas de jogar no Battlegrounds: no primeiro modo, os times são formados de forma aleatória – os jogadores escolhem entrar na partida e esperam os times se formar, e jogam de forma que possam cooperar entre si para derrotar o time inimigo. A segunda forma é uma experiência mais completa de time, onde um grupo de jogadores mais organizados e com uma estratégia mais completa se une e joga contra outros grupos de nível similar para ganhar pontos de conquista, o que permite

ao time vencedor comprar itens mais interessantes, além de poder ter um ranking na maior pontuação dos Battlegrounds da região.

Outland e Northrend eram, na época, as duas regiões que jogadores iam para completar missões, ganhar nível, buscar itens com seus personagens. Cada uma dessas regiões possui missões específicas para cada nível de jogador e é onde os jogadores vão com seus grupos, ou sozinhos, evoluir seus personagens. Essas regiões também possuem os portais de chegada, que as Capitals enviam personagens para facilitar a viagem e economizar tempo. Os jogadores aqui podem também desbravar a região em busca de itens ou monstros que não são comumente usados na linha de evolução dos personagens.

Por fim, a Arena é a região onde jogadores vão para se digladiar em grupos, porém, esses grupos não precisam ser compostos por jogadores da mesma facção ou *guild*, são simples confrontos onde somente um time sai vencedor e isso alimenta um *ranking* de *Arena*. Aqui, os jogadores se divertem competindo contra outros jogadores, obtendo o *rank* mais alto e se mostrando e temidos quando entram nessas competições.

Em cada uma das regiões foi possível notar que possuem alguma semelhança com os tipos de jogadores de Bartle. Apesar de não poder falar que cada região favorece 100% de um tipo específico, cada região favorece mais o surgimento de um tipo que outras. Para realizar o experimento, foram definidas as seguintes premissas: As regiões Arena e Capitals são as regiões que possuem as características que tendem a agradar aos *Socializers*, por serem as regiões mais comuns para socializar com outros jogadores, ou até mesmo para interpretar os seus personagens.

A região *Arena* pode indicar uma área de batalhas PvP, porém, por ser uma região que só se entra com um grupo de jogadores, foi classificada como *Socializers*.

A região Battlegrounds foi classificada para os Killers, uma vez que é a região específica para o confronto entre jogadores.

Por fim, Northrend e Outland foram classificadas como regiões onde os Explorers e Achievers se concentram, por ser as regiões onde as RAIDs, as missões de cada classe, se encontram, e onde os jogadores vão para coletar os itens raros, evoluir seus avatares, explorar o mapa e suas rotas.

Ao definir essas premissas, foram realizados alguns testes iniciais para entender melhor a base de dados, e verificar quais as informações que seriam utilizadas para identificação dos tipos de jogadores. Inicialmente foi escolhido um grupo inicial, para manter o controle das informações e manter as análises menos complexas para o primeiro momento. Isso pode ser visto na figura 18, no passo (2).

Em um segundo momento, após as variáveis terem sido definidas, o trabalho seguiu a metodologia apresentada no início, seguindo os passos (3) a (6) (Figura 18).

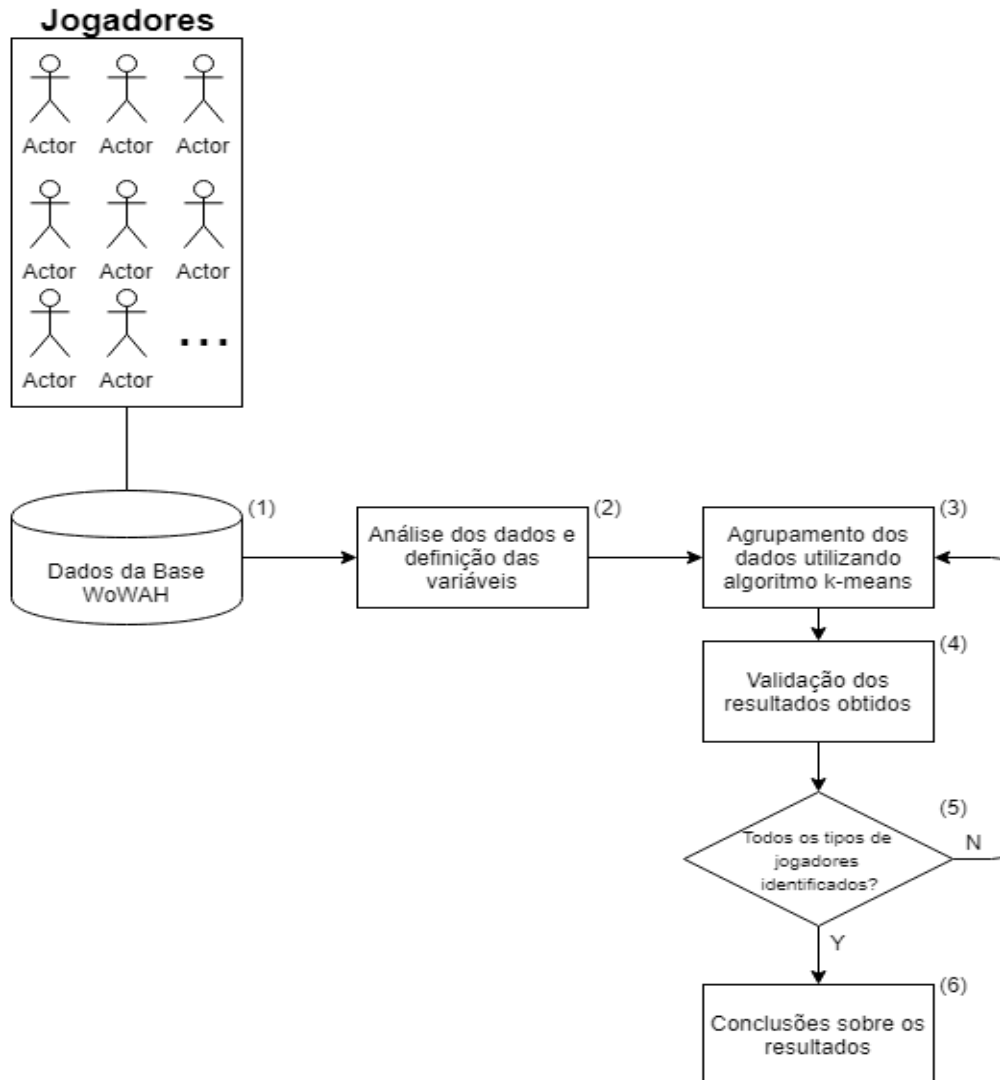


Figura 18 – Metodologia usada para identificar os tipos de jogadores, usando k-means
Fonte: Criação do autor

Os dados dos jogadores escolhidos foram os personagens, as regiões que visitam, a quantidade de horas jogadas nessas regiões e a quantidade de vezes que essas regiões são visitadas.

Com essas informações, é realizado um primeiro agrupamento utilizando k-means (3) e, logo em seguida os resultados são validados (4). Será mostrado que, com apenas um agrupamento não foi possível descobrir os quatro tipos de jogadores e, para isso, um novo agrupamento foi necessário, por isso há o passo (5), de verificar se todos os tipos são identificados. Por fim, é possível descrever as conclusões e discutir os resultados obtidos (6).

No capítulo 2, foi mostrado que não há processos bem definidos de *game analytics* e *game data mining*. Com isso, o presente trabalho definiu dois frameworks para trabalhar com esses dois termos, baseados nas metodologias apresentadas em Drachen, El-Nasr e Canossa (2012).

A figura 19 apresenta como o *Game Analytics* foi utilizado no trabalho. Primeiramente, foi preciso definir a telemetria do jogo, que é dividida em duas partes: a coleta dos dados, e os dados que são coletados.

Neste trabalho não foi realizada nem a coleta nem o armazenamento, pois fora feito anteriormente em Lee et al. (2011), e esses dados coletados foram utilizados diretamente como a fonte do *Game Metrics*. Após definir a fonte, o passo seguinte foi definir as informações relevantes, que foram utilizadas para realizar o *Game Data Mining*.

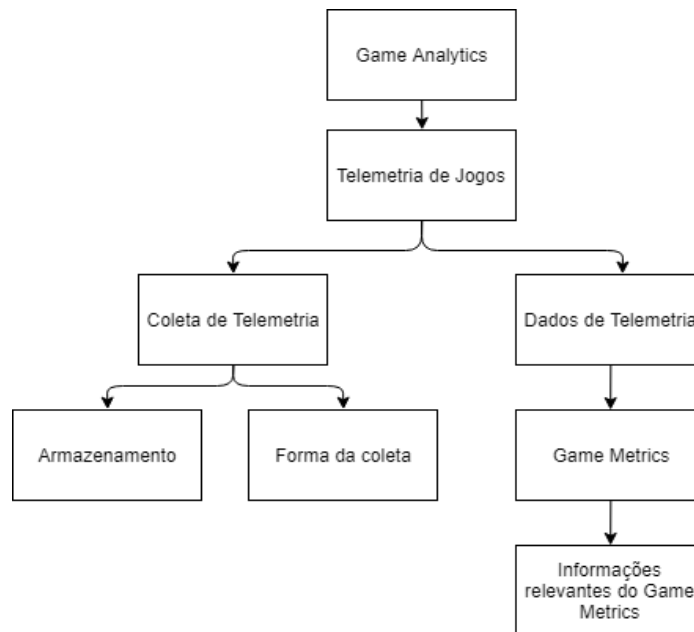


Figura 19 – Framework de *Game Analytics*

Fonte: Criação do autor

A figura 20 tem o framework utilizado para realizar os passos do *Game Data Mining*. Primeiramente, foi necessário separar as informações que seriam utilizadas para realizar o *data mining*. Em posse dessas informações, foi preciso definir a técnica a ser utilizada, que no caso foi o *clustering*, através do k-means. Após isso, os dados foram normalizados e ajustados para que fossem válidos e de fácil utilização ao utilizar o algoritmo de agrupamento. O próximo passo foi utilizar a técnica para realizar o *data mining*. Por fim, os resultados foram avaliados e as conclusões anotadas.

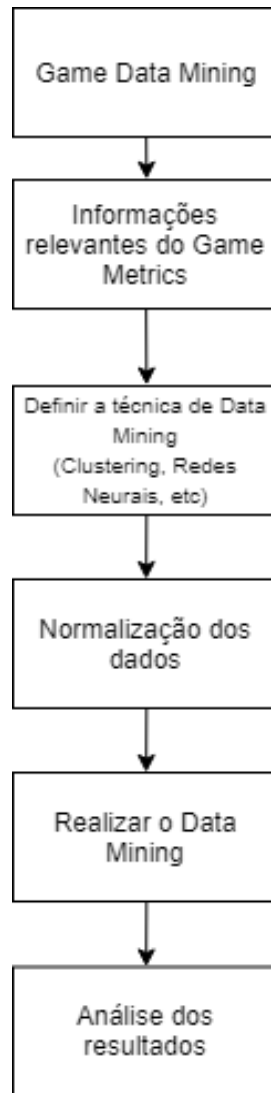


Figura 20 – Framework de *Game Data Mining*
Fonte: Criação do autor

A seguir serão mostrados os resultados de ambos os experimentos.

4.3 APRESENTAÇÃO DOS RESULTADOS DO EXPERIMENTO INICIAL

O objetivo do experimento era capturar os dados do maior grupo (*guild*), pois assim é possível obter um grupo controlado de jogadores. O primeiro passo foi identificar a maior *guild* do jogo, como mostra a figura 21. A *guild* com maior número de ocorrências foi a 103.

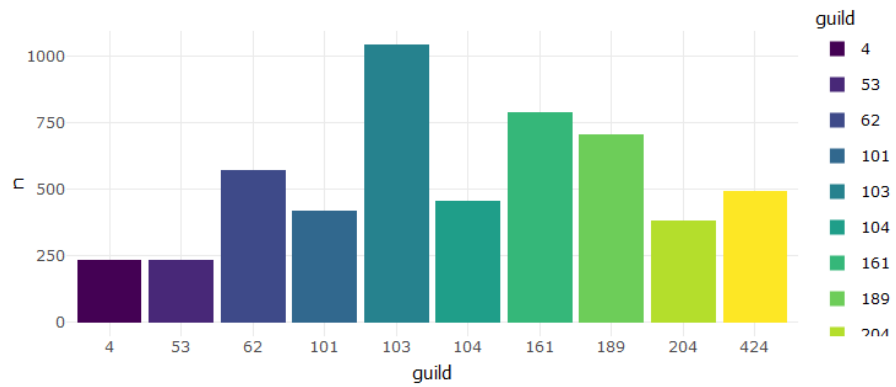


Figura 21 – Número de jogadores em cada uma das top 10 guilds
 Fonte: Criação do autor

Uma vez identificada essa *guild*, foi feita uma visualização da distribuição das classes desses jogadores, e um mapa de calor das classes e raças dos Avatares, como mostrado nas figuras 22 e 23.

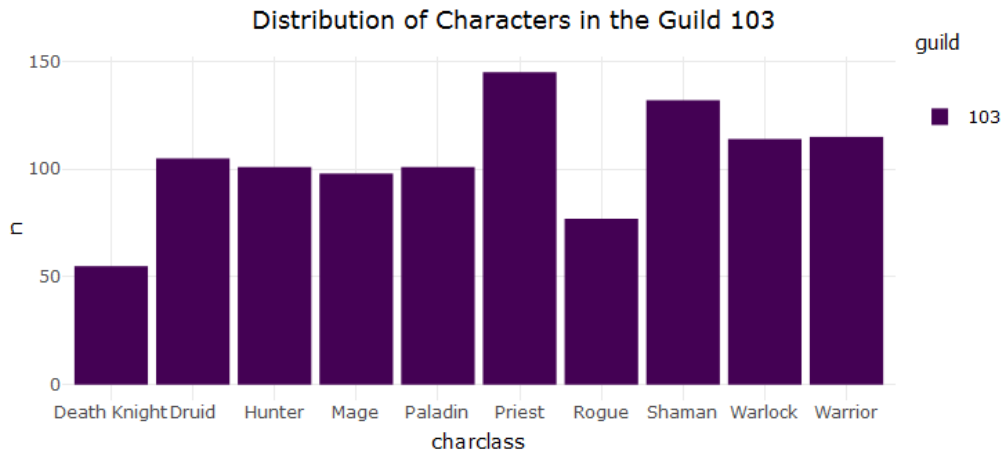


Figura 22 – Distribuição de Classes de Avatares na *Guild* 103
 Fonte: Criação do autor

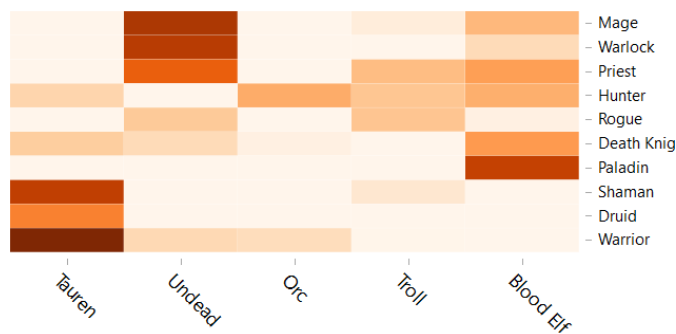


Figura 23 – Mapa de calor das classes e raças dos Avatares da *Guild* 103
 Fonte: Criação do autor

Na figura 23, quanto mais escura a cor, maior a incidência de resultados com a combinação Raça x Classe e, com esse resultado é encontrada a composição da *guild* 103: A classe de maior ocorrência é a *Priest*, porém a que possui maior combinação é a de *Warrior Tauren*.

Essas são duas classes que combinam com o estilo de uma *guild*, uma vez que, para completar missões juntos, sempre há a necessidade de se ter um personagem que seja capaz de sofrer muito dano (*Warriors*) e outro capaz de recuperar vida (*Priest*).

Esse tipo de informação pode ser útil para ajudar a identificar os perfis dos jogadores, que tendem a ser *Socializer* e *Killer*, pois cada tipo de jogador tende a escolher um personagem que o complete. Em geral, *Killers* tendem a escolher raças que são autossuficientes e aguentam e causam mais danos sem depender de outros (como Paladinos e Guerreiros). *Socializers* tendem a escolher a melhor classe para o grupo, e geralmente esses jogadores escolhem ser do tipo suporte, que são classes que ajudam a melhorar o desempenho de outros jogadores (como Magos e Druidas).

A figura 24 mostra a distribuição dos jogadores ativos durante os meses, e seus níveis durante o período. A maioria dos jogadores da *guild* 103 já começa com avatares entre os níveis mais altos do servidor (70) e, com o lançamento da expansão *Warth of the Lich King* (WotLK) em outubro de 2008, observamos os personagens saindo do nível 70 e subindo para o nível 80.

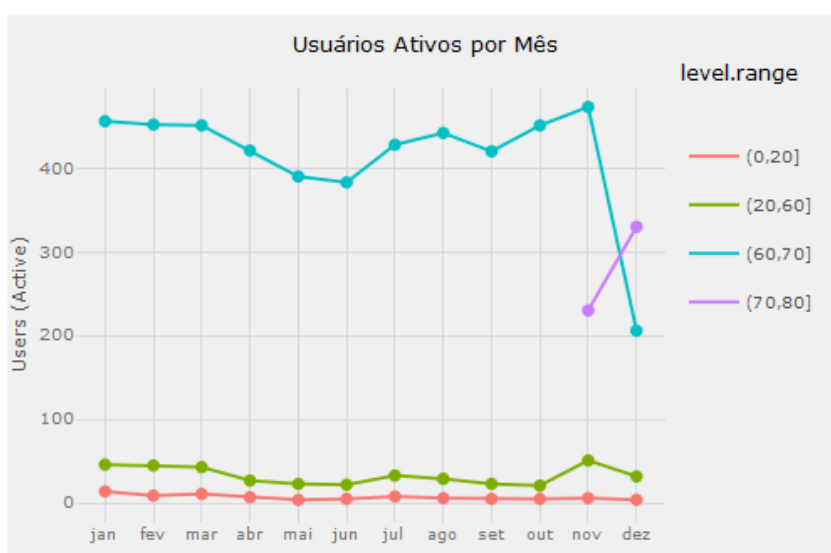


Figura 24 – Total de usuários ativos por mês, no ano de 2008

Fonte: Criação do autor

Com esse gráfico, pode-se começar a identificar os Achievers, pois em menos de quinze dias do lançamento da expansão WotLK, já são vistos os jogadores no nível máximo ou próximo dele.

Na figura 25, é mostrada a quantidade de horas jogadas de cada jogador da *guild* 103, através dos dias, e pode-se concluir que a maioria dos jogadores da *guild* 103 joga de 0 a 7 horas por dia. Também é interessante notar que há três períodos em que não há medição, por conta da manutenção dos servidores, deixando os jogadores impossibilitados de jogar no período.

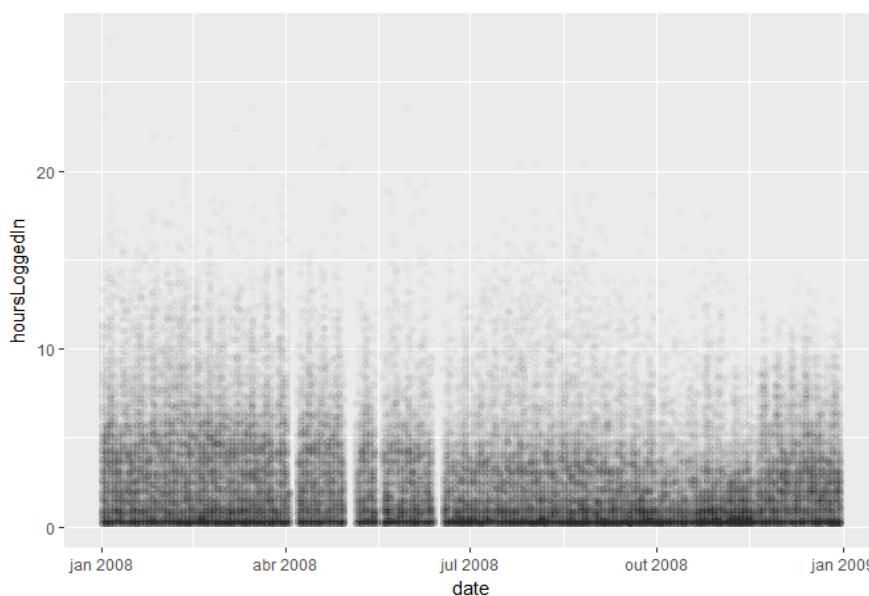


Figura 25 – Distribuição de horas jogadas de cada jogador da *guild* 103

Fonte: Criação do autor

Esse tipo de informação é útil para auxiliar na identificação dos *Explorers* e *Achievers*. *Explorers* querem conhecer detalhadamente cada parte do mundo, o que pode demorar muito tempo no jogo, enquanto *Achievers* querem obter o maior número de itens ou colecionáveis do jogo, que muitas vezes dependem da taxa de recompensa de cada monstro matado, fazendo com que esses dois tipos de jogadores gastem uma maior quantidade de horas jogando. Itens raros, por exemplo, só são possíveis de coletar após completar missões especiais, que podem demorar algumas horas, e nem sempre é garantido, o que requer repetição da missão até receber o item desejado.

As figuras 26 e 27 mostram uma visão de quantas pessoas jogam no mesmo horário. Nos dois semestres, os jogadores da *guild* 103 se concentram no horário das 20 horas em diante. Percebe-se que a intensidade nesse horário é maior nos meses de fevereiro a março; após esse período, é encontrado um outro período em que se mantém uma média de quase 100 jogadores online no horário, até o mês de novembro, que é quando a expansão WotLK foi lançada, onde também é observado um aumento na quantidade de jogadores online ao mesmo tempo, como é

visto nas figuras 25 e 26. Em uma análise um pouco mais detalhada, poderá ser feita uma correlação dessas informações com a quantidade de horas jogadas de cada jogador, para identificar a real motivação dos jogadores e definir seu arquétipo.

Ainda é possível perceber as mesmas datas que os jogadores ficaram sem acessar o jogo nos meses de abril, maio e junho de 2008.

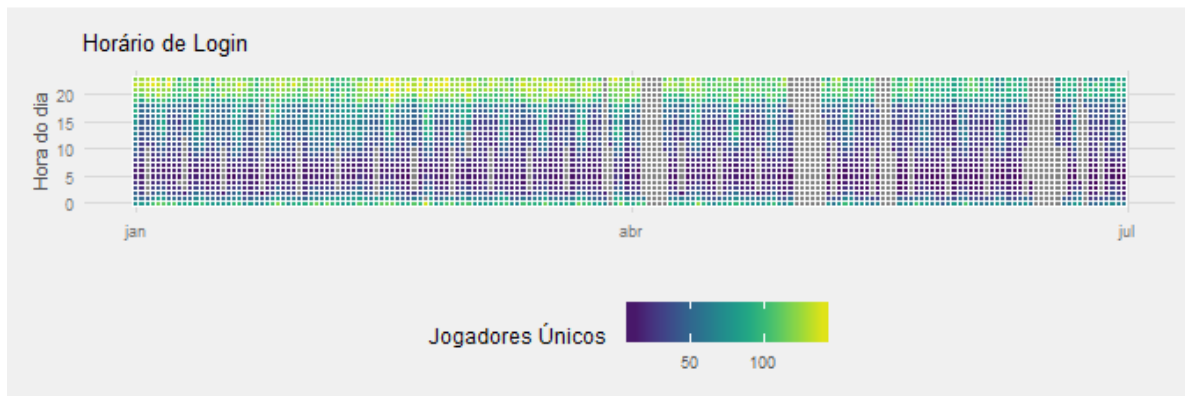


Figura 26 – Horário de *Login* dos jogadores da *guild* 103, no primeiro semestre de 2008
Fonte: Criação do autor

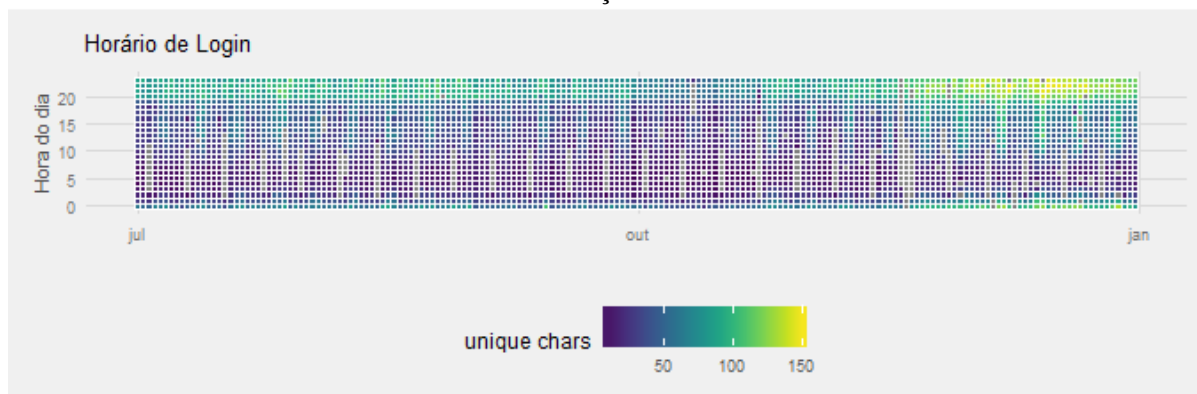


Figura 27 – Horário de *Login* dos jogadores da *guild* 103, no segundo semestre de 2008
Fonte: Criação do autor

Na figura 28 é mostrada uma visão de quanto tempo uma classe leva para chegar até o nível máximo. Na figura, cada linha colorida representa a quantidade de dias que um jogador levou para chegar ao nível máximo (80), durante o período analisado. Há uma nítida diferença do tempo que precisa para evoluir cada tipo de classe de personagens. *Rogues*, *Shaman*, *Warlocks* e *Druids* são os que evoluem mais rápido nessa *guild*, enquanto *Warriors* e *Priests* são os que demoram mais. É importante notar que os *Death Knight* foram lançados apenas na expansão WotLK, disponibilizados no nível 55 e com a restrição de um personagem dessa classe por conta.

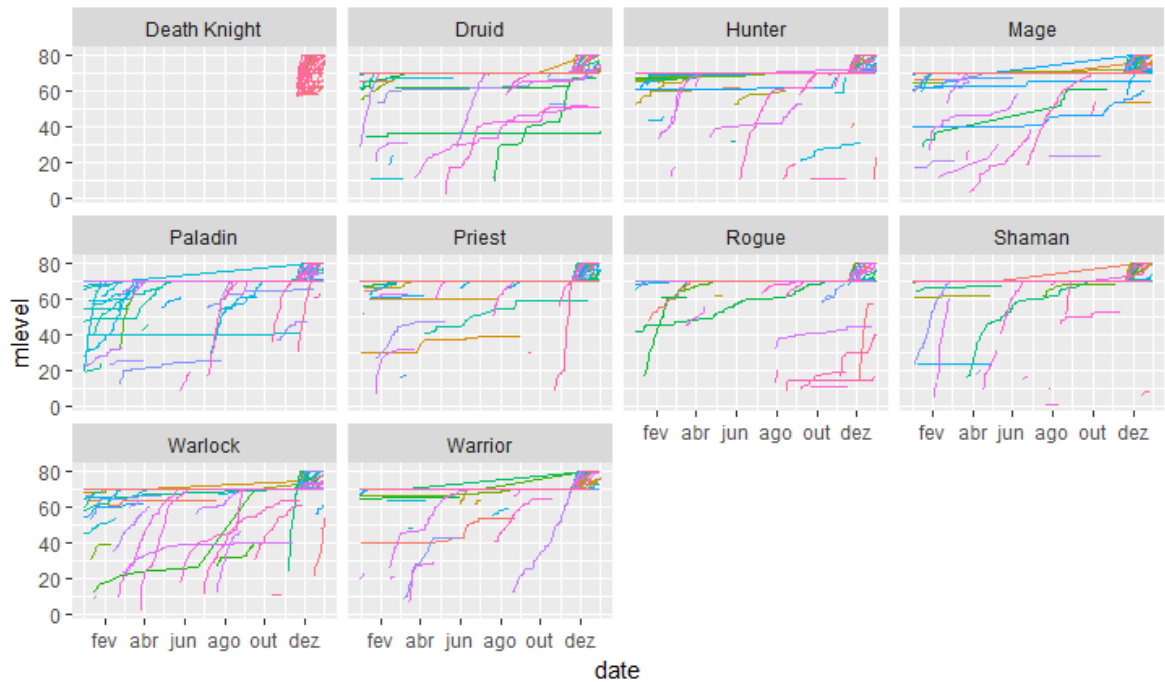


Figura 28 – Tempo para atingir o nível máximo de cada classe
Fonte: Criação do autor

Uma investigação que pode ser feita é se os jogadores que jogam juntos utilizam da *guild* para conseguirem subir de nível mais rápido. A figura 29 mostra as regiões mais acessadas, de acordo com os níveis de cada jogador.

Jogadores de níveis iniciais, ou mesmo mais baixos do que a região sugerida pelo jogo, podem se juntar a jogadores de níveis altos e ganhar experiência mais rápido, uma vez que a experiência é compartilhada com jogadores da mesma *party* – *party*, grupo formado entre jogadores que escolhem jogar juntos, sendo da mesma *guild* ou não.

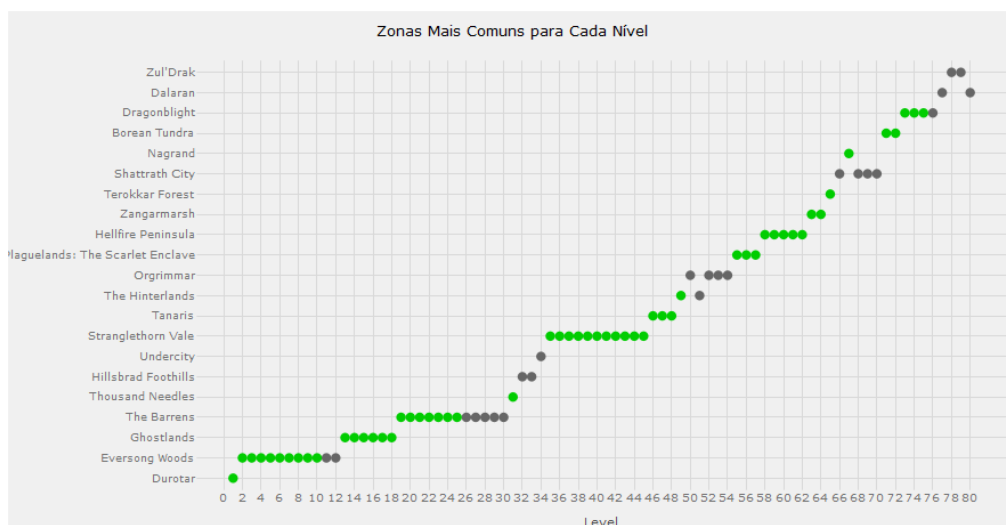


Figura 29 – Zonas mais acessadas entre jogadores de cada nível
Fonte: Criação do autor

As figuras 30, 31, 32 e 33 mostram o deslocamento dos avatares entre os níveis 70 e 80 dentro do jogo. A figura 30 mostra o deslocamento global desses jogadores.

Os jogadores foram escolhidos no nível mais alto, para ver a movimentação deles nas áreas de Arena (figura 31), onde se concentram os duelos PvP. Só se tem acesso a essa região nesse nível com a intenção de encontrarmos os jogadores do tipo Killer, e pode-se notar que a movimentação para essas cidades é grande.

Outra região interessante para se olhar é a região Battlegrounds (figura 32), onde as *guilds* entram para se enfrentar, porém a concentração não é grande dentro dos jogadores desse nível, ao comparar com as outras áreas.

A maior concentração desses jogadores é nas cidades ditas Capitals (figura 33). Nessas cidades, os jogadores podem encontrar atrativos para ficarem, pois possuem locais para compra e venda de itens, casa de leilões, bancos, entre outras atividades que chamam a atenção para jogadores desse nível. São boas localizações e, para entender quais jogadores procuram essas cidades, um jogador do tipo *Socializer*, por exemplo, pode ficar ali pois poderá socializar com outros jogadores, devido à alta concentração de jogadores. Já os *Achievers* podem comprar itens raros nas casas de leilão e os *Explorers* podem procurar por caminhos mais rápidos para regiões diferentes com portais.

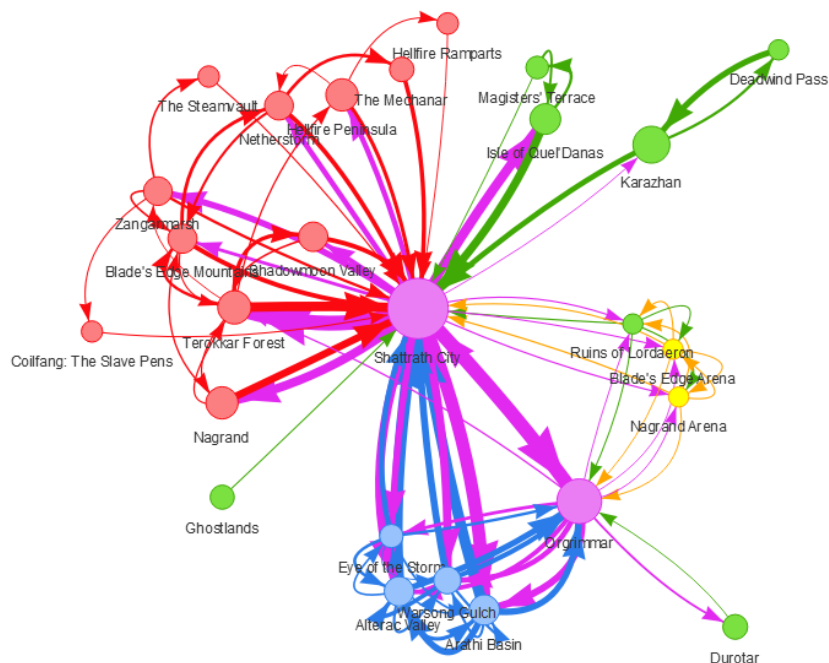


Figura 30 – Deslocamento global dos jogadores de nível entre 70 e 80 da *guild* 103

Fonte: Criação do autor

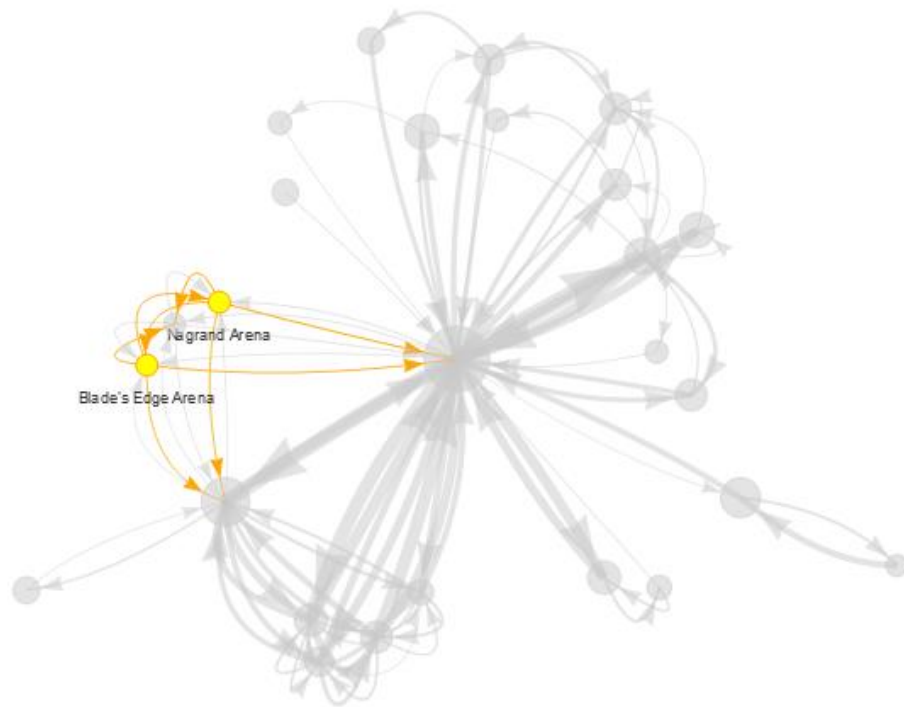


Figura 31 – Deslocamento nas regiões Arena
 Fonte: Criação do autor

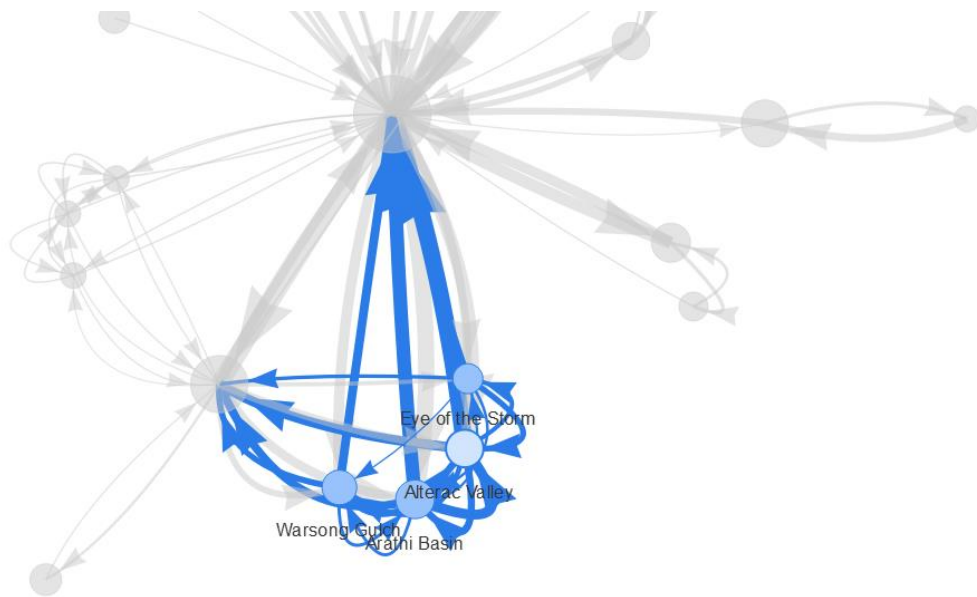


Figura 32 – Deslocamento dos jogadores na região *Battlegrounds*
 Fonte: Criação do autor

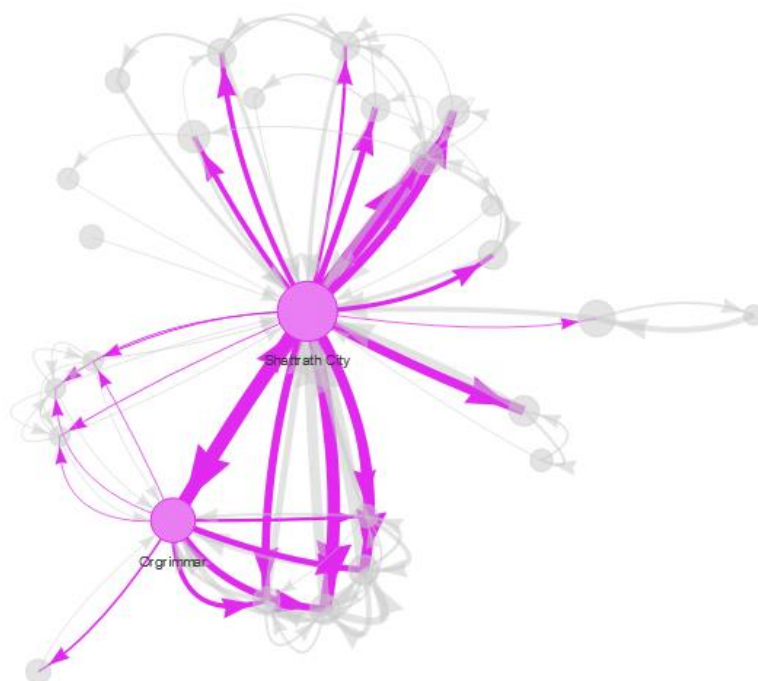


Figura 33 – Deslocamentos dos jogadores a partir das Capitais
Fonte: Criação do autor

4.4 AGRUPAMENTO E CLASSIFICAÇÃO DOS JOGADORES

Após a realização dos experimentos iniciais, foi observado que, para realizar a classificação dos jogadores, as melhores variáveis disponíveis seriam a quantidade de jogadores únicos e a quantidade de horas jogadas, em cada uma das regiões do jogo: Capital, Northrend, Battlegrounds, Outland e Arena.

Nas figuras 34 e 35, cada círculo representa um jogador distinto da guild 103. Os quatro grupos podem ser identificados facilmente e são distintos entre si, de acordo com a quantidade de horas jogadas. Na figura 34, o número de identificação dos jogadores é definido pela variável char (eixo x) e, na figura 35, cada região precisou passar por uma normalização, e o resultado pode ser conferido a seguir: 1 – Capital, 2 – Battlegrounds, 3 – Northrend, 4 – Outland, 5 – Arena.

4.4.1 Agrupamento dos jogadores da guild 103

Com os quatro tipos de jogadores separados em cada uma das regiões, a base de dados da Guild 103 foi submetida ao algoritmo k-means. O número de *clusters* definido foi o dos tipos de jogadores de Bartle (4), e os resultados podem ser vistos nas figuras 34 e 35.

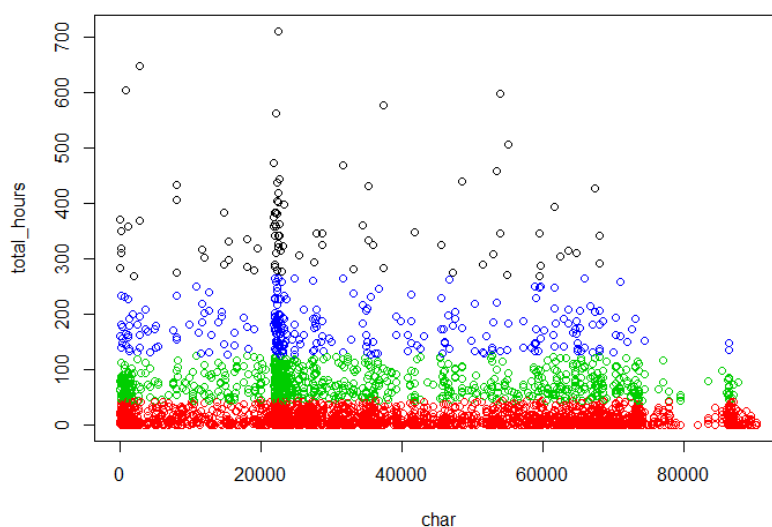


Figura 34 – Grupos de jogadores, a partir da quantidade de horas jogadas
Fonte: Criação do autor

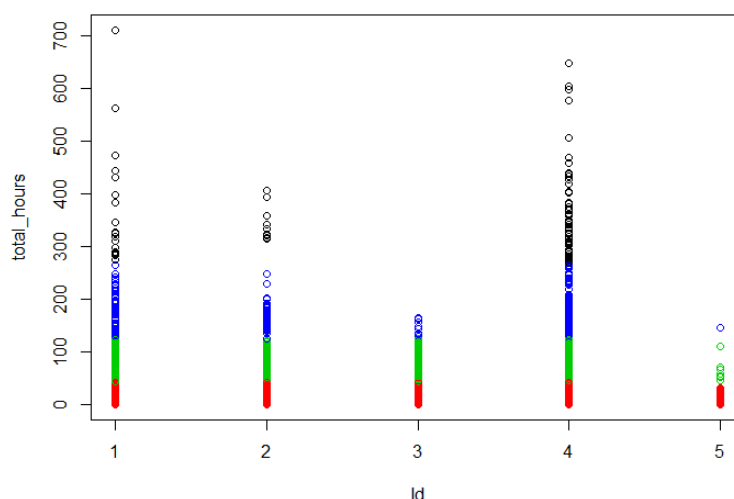


Figura 35 – Quantidade de horas jogadas em cada uma das regiões
Fonte: Criação do autor

Pode ser observado que essa guild passa muito mais horas jogando nas regiões Capital e Outland que os tipos de jogadores Killers, que gastam suas horas na região 5, e é uma população pequena nessa *guild*. Com essa observação, pode-se concluir que os jogadores *Socializers* e *Killers* são facilmente identificados, enquanto que, para descobrir os jogadores *Achievers* e *Explorers* será necessária uma análise detalhada das regiões Outland e Northrend.

Para identificar os jogadores *Achievers* e *Explorers*, foi escolhido contar quantas vezes os jogadores visitavam as regiões 3 e 4, e quantas horas eram jogadas nessas regiões. Novamente o algoritmo k-means fora aplicado, porém, como era desejado encontrar apenas dois grupos de jogadores, o k escolhido foi dois. O resultado pode ser visto na figura 36, onde é possível identificar que foram obtidos dois grupos distintos: um que joga até 150 horas e que

retorna muitas vezes nessas regiões, e um segundo grupo com mais de 150 horas jogadas, com uma frequência de retorno concentrada entre 10 e 20 vezes.

O primeiro grupo (preto) pode ser classificado como o dos Achievers, uma vez que interagem com as regiões apenas para alcançar seus objetivos e, por mais que retornem muitas vezes, a quantidade de horas gastas mostra que a tendência é ir para essas regiões, coletar as recompensas e escolher outros objetivos.

O segundo grupo (vermelho), foi definido como o dos Explorers, devido à quantidade de horas jogadas e o pouco retorno nessas regiões. Uma vez que a região é explorada, esse tipo de jogador prefere ir para outra região e começar outros objetivos.

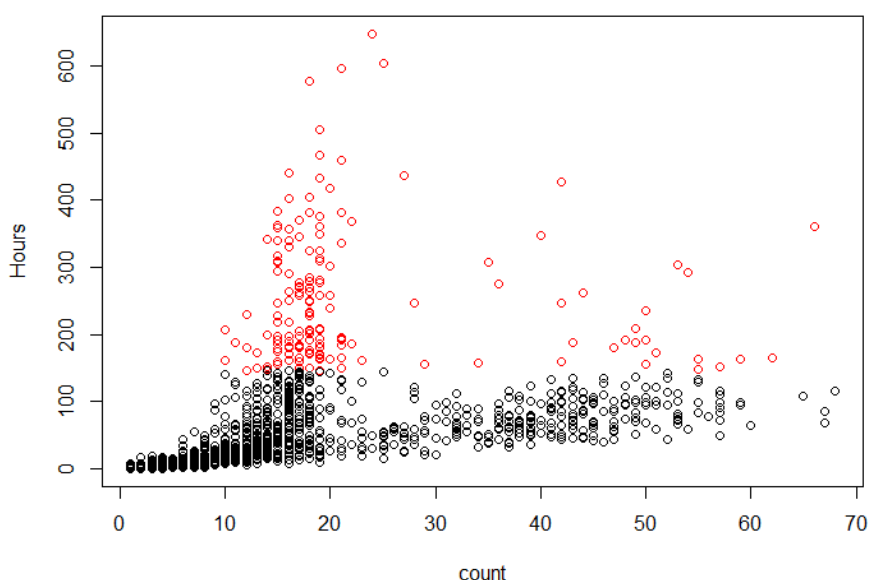


Figura 36 – Agrupamento de jogadores das regiões Northrend e Outland, pela quantidade de horas jogadas e quantidade de visitas nessas regiões

Fonte: Criação do autor

4.4.2 Agrupamento com todos os jogadores

Logo após realizar o experimento com um grupo controlado, foi utilizado o mesmo procedimento, detalhado anteriormente na metodologia, com a base de dados de um ano inteiro de captura de telemetria dos jogadores.

Apesar da quantidade de jogadores ser maior, o resultado obtido é similar ao experimento reduzido, onde pode-se encontrar a maioria dos jogadores jogando mais tempo em regiões que tendem a ser para *Socializers*, *Explorers* e *Achievers*, enquanto que a região em tese mais propícia para Killers é a que possui a menor quantidade de horas jogadas.

O resultado é mostrado nas figuras 37, 38 e 39. Na figura 37, pode-se notar que os grupos são divididos pela quantidade de horas jogadas com números bem próximos ao do experimento anterior, porém com um jogador atingindo um total de mais de mil horas jogadas no ano. Apesar desse valor, não se compromete os resultados obtidos na divisão dos grupos.

Na figura 38, mostra-se como os jogadores se dividem em quantidade de horas jogadas, em cada uma das cinco regiões. Similar ao grupo controlado, os jogadores não jogam muitas horas na região Arena (5), e se concentram nos duos de regiões Capital/Battlegrounds e Outland/Northrend.

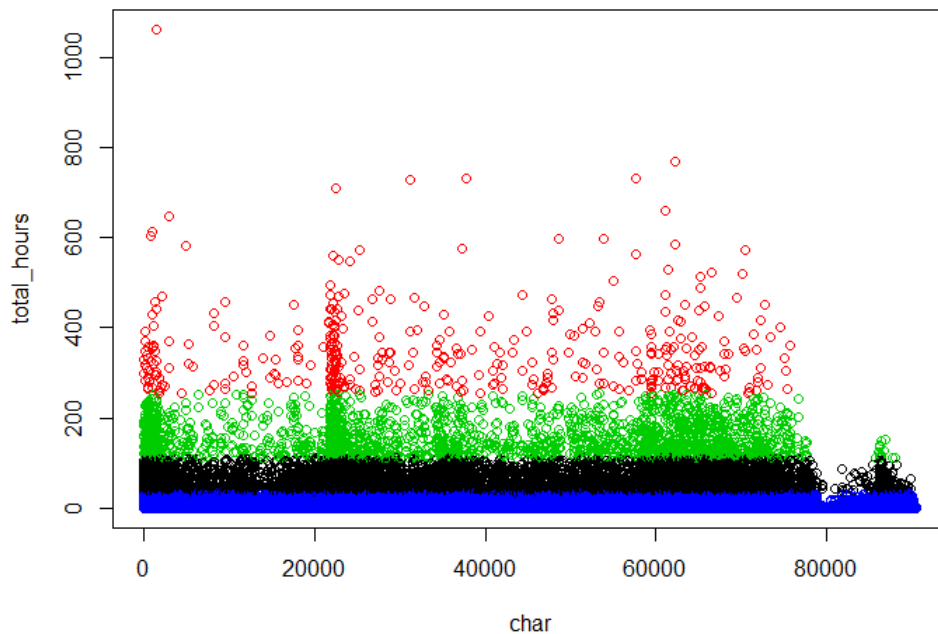


Figura 37 – Grupos de todos os jogadores
Fonte: Criação do autor

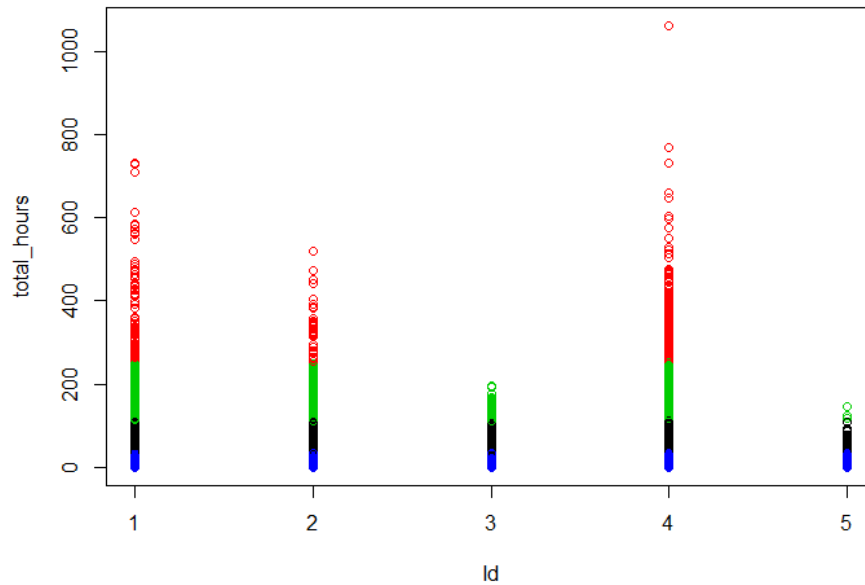


Figura 38 – Jogadores agrupados pela quantidade de horas jogadas em cada região
Fonte: Criação do autor

Como no experimento anterior, foi necessário agrupar novamente os jogadores para identificar os tipos *Explorers* e *Achievers*. A figura 39 mostra como os jogadores se dividem em quantidade de horas jogadas e na quantidade de vezes que jogam em cada uma dessas duas regiões. Os jogadores podem ser identificados como *Achievers* dentro do grupo vermelho, por jogarem menos horas e não voltarem menos vezes nessas regiões, o que mostra que, ao atingirem seus objetivos, preferem jogar em outras regiões. Os do grupo preto podem ser classificados como *Explorers*, devido à grande quantidade de horas gastas nessas regiões e a quantidade de vezes que voltam à região. Isso é mostrado na figura 39 abaixo.

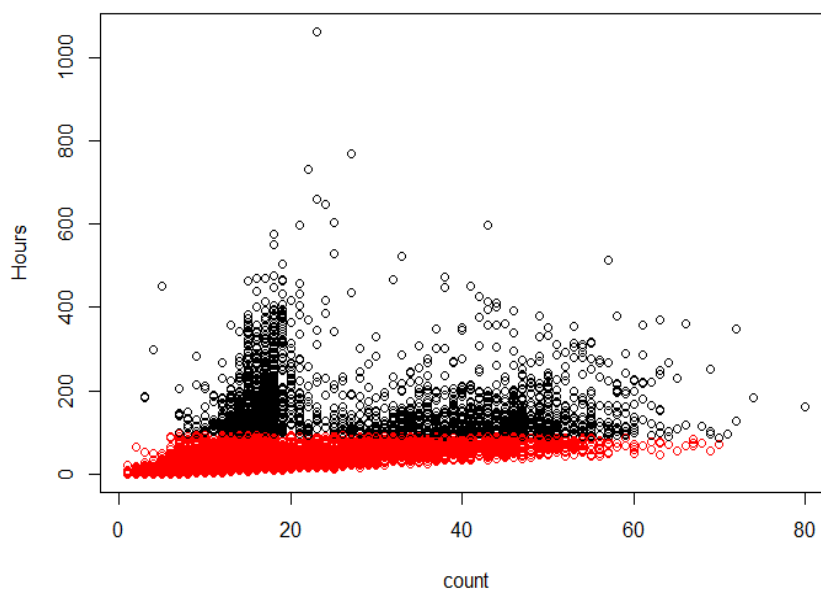


Figura 39 – Número de vezes que as regiões Northrend e Outland foram visitadas, quantidade de horas jogadas e visitas nessas regiões
Fonte: Criação do autor

Com o agrupamento, foi possível identificar a quantidade de jogadores por *cluster* e por consequência, foi possível identificar também os percentuais dos tipos de jogadores no servidor durante o período, como visto nas tabelas 3 e 4.

Na tabela 3, temos como os jogadores são distribuídos dentro de cada *cluster* do k-means. Seus valores são da quantidade absoluta de jogadores em cada um desses grupos. O maior deles é o 3, enquanto o menor é o 1.

Tabela 3 – Distribuição de jogadores por *cluster*

Regiões	Clusters			
	1	2	3	4
Capital	100	550	25513	2045
Battleground	35	280	6068	996
Northrend	0	93	2800	912
Outland	260	1087	8177	2546
Arena	0	4	2729	52
Número de jogadores por <i>cluster</i>	395	2014	45287	6551

Fonte: Criação do autor

Com o auxílio desses valores obtidos, foi possível calcular a distribuição dos jogadores nas regiões, de acordo com o seu tipo de jogador. *Socializer* representa 66% dos jogadores nessa metodologia, enquanto *Achiever* é o menor grupo, com 4% de representatividade. Isso mostra que, no ano de 2008, os jogadores passaram a maior parte do tempo em regiões em que precisaria interpretar melhor seu personagem, para utilizar

todas as suas habilidades de forma efetiva, do que em qualquer outra região. Em segundo lugar ficaram os jogadores que preferiram gastar seu tempo andando pelas regiões e descobrindo seus segredos e, praticamente empatados, os que preferem participar de combates PvP e os que preferem coletar itens ou subir de nível mais rápido.

Tabela 4 – Distribuição de jogadores por arquétipo

Socializer	Killer	Achiever	Explorer
35587	2785	1921	13954
66%	5%	4%	26%

Fonte: Criação do autor

Esses padrões podem ser considerados, devido a 2008 ter sido um período sem muitas atualizações dos servidores de WoW, fazendo com que seus jogadores ficassem um período longo sem novas missões ou novos itens a serem coletados. Após 22 meses de sua última grande atualização, em novembro de 2008 a expansão *Wrath of The Lich King* é lançada e faz com que os jogadores possam evoluir mais seus avatares (o nível máximo a ser atingido passa de 70 para 80). Uma nova classe é liberada para os jogadores e novas áreas com missões e itens para serem coletados (WOWWIKI, 2018)

5 CONCLUSÃO E TRABALHOS FUTUROS

Nesta pesquisa, foi possível investigar dados de jogadores de um jogo MMORPG, organizá-los e classificá-los em grupos e analisar sua aderência a propostas de arquétipos bem estabelecidas na literatura. Isso possibilitou encontrar os padrões de jogabilidade que interessam aos desenvolvedores dos jogos, para que possam fazer alterações e satisfazer o maior número de jogadores possível, fazendo assim com que permaneçam em seu ambiente por mais tempo.

Na revisão bibliográfica, foram encontradas situações similares com a proposta deste trabalho (DRACHEN; CANOSSA; YANNANAKIS,2009), (DRACHEN et al, 2012), porém, em ambos os casos uma nova taxonomia referente aos jogos estudados foi criada e, com essa nova classificação os jogadores foram analisados.

Mesmo com a falta de informações mais detalhadas no grupo de dados, os quatro tipos de jogadores de Bartle puderam ser identificados. Há que se ressaltar que *Socializers* poderiam ser mais facilmente identificados se na base de dados tivesse a quantidade de mensagens trocadas durante o período, ou a quantidade de amigos *in-game* que cada jogador possui, apenas para ficar com esse exemplo.

Outro ponto importante a ser destacado foi a importância que o tipo de jogador Explorer tem dentro desses jogos. Hoje em dia, os jogos estão cada vez mais limitando quaisquer ações de jogadores que não sejam as previamente programadas, fazendo com que esses jogadores busquem informações fora do jogo, como em fóruns de discussão, para buscar qualquer que seja o ponto que possa ser explorado. Esse tipo de comportamento dificulta, e muito, a sua identificação através do seu gameplay e qualquer tipo de possibilidade dos desenvolvedores e administradores de servidores de buscar e implementar as melhorias que atraiam esse tipo de jogadores ao seu jogo.

Acerca de *game analytics*, foi identificado que não há frameworks bem definidos e por isso foi proposto no capítulo 4 dois frameworks que foram utilizados no trabalho, baseados nas metodologias apresentadas em Drachen, A.; El-Nasr, M. S.; Canossa, A (2012), que ficam para serem analisados em um trabalho futuro em caso de melhorias para ambos frameworks.

Outra dificuldade identificada foi a interação entre cada um dos tipos de jogadores, como se comporta cada um dentro do jogo e quais são seus reais objetivos. Tanto a região Arena quanto Battlegrounds possuem rankings, e por isso podem muito bem atrair a atenção dos jogadores do tipo Achiever. Essa interação possui um limiar muito tênue, que não foi levado

em consideração no presente trabalho por não ser o escopo do trabalho, mas pode ser um objeto de estudo futuro a respeito do objetivo de cada jogador no jogo.

Apesar das dificuldades encontradas, o método de agrupamento de dados foi bastante eficiente na composição dos grupos e, com base nas premissas adotadas, os tipos de jogadores se mostraram bem distintos na sua forma de jogar e locais de preferência.

Socializers têm a preferência de interpretar seus personagens com a maior fidelidade possível, e interagir com outros jogadores. As regiões Capital e Battlegrounds apresentam essas características e onde os jogadores mais se concentraram no ano dessa amostra. Jogadores Killer querem se sobressair e se impor aos outros jogadores. Na região Arena eles conseguem fazer isso e mostrar quem é o melhor jogador daquele lugar. Explorers vão em busca do campo aberto do mapa onde podem encontrar os lugares mais escondidos, ou até mesmo uma paisagem bonita do jogo para ter uma foto do jogo que ninguém possui, enquanto os Achievers vão em busca das regiões com maior possibilidade de obter aquele item que tem uma probabilidade baixa de encontrar, por isso as regiões Northrend e Outland os satisfazem tanto.

Essas possibilidades proporcionam à empresa desenvolvedora do jogo, por exemplo, realizar ações de marketing direcionado a um público específico, melhorar alguma mecânica de jogo, criar missões novas que possam trazer jogadores antigos inativos, entre outras ações, para fazer com que seu jogo continue relevante no mercado.

Essas informações são o ativo mais importante da empresa, uma vez que não as divulga e nem compartilha com ninguém de fora, por isso, conhecer seus jogadores cada vez mais pode fazer com que eles influenciem cada vez mais novos jogadores, ou mesmo aqueles que não jogam mais, para o seu mundo virtual.

Nesta pesquisa, foram apresentados os fundamentos de *Game Analytics*, Telemetria de dados e *Game Data Mining*. É fundamental entender que as ferramentas auxiliam na tomada de decisões e, até o presente momento, seres humanos são parte fundamental para o preparo e entendimento dos dados obtidos.

Não foi objetivo do presente trabalho ser uma fonte detalhada para os princípios de *Data Mining*, e sim de utilizar algumas de suas técnicas e métodos para a descoberta de padrões de dados.

Por fim, devido ao aumento na procura por melhores resultados e desempenho no ramo de jogos, há disponível no mercado alguns programas desenvolvidos especificamente para *Game Data Mining* e *Game Analytics*, como os que podem ser encontrados nos web sites a seguir: www.gameanalytics.com, www.playtomic.com, www.honeytracks.com, www.kontagent.com (DRACHEN et al., 2013).

Como trabalhos futuros fica a sugestão de realizar a análise de movimentos entre regiões e o tempo para subir de nível dos avatares, para que seja verificado se há correlação entre essas informações, a fim de que possam ser mais bem identificados os jogadores *Explorers* e *Achievers*, obtendo um método de classificação que vá além de apenas analisar a quantidade de horas jogadas. Outros possíveis trabalhos futuros apontam para a replicação do experimento em outras bases de jogadores, bem como a análise à luz dos oito arquétipos de Bartle, segundo a revisão do modelo feita por esse, ou ainda outros arquétipos, como os de Marzewsky, por exemplo.

Ainda como proposta de seguir a pesquisa adiante pode-se calcular os jogadores através do centroide de cada grupo, ou analisar com um outro algoritmo (k-medoid) e observar se os resultados obtidos serão muito diferentes dos encontrados nesta.

REFERÊNCIAS

ANDREASEN, E.; DOWNEY, B. The mud personality test. *The MUD Companion*, v. 1, p. 33-35, 2001.

MARCZEWSKI, ANDRZEJ. User types & player types in gamification #GWC14. In: GAMIFICATION WORLD CONGRESS, 2014. Madrid. *Anais...* Spain. Disponível em: <<https://www.youtube.com/watch?v=R6C9d8uGKV0>>. Acesso em: 16 out. 2017.

BARROS, V. P.; NOTARGIACOMO, P. Big data analytics in cloud gaming: players' patterns recognition using artificial neural networks. In: IEEE INTERNATIONAL CONFERENCE ON BIG DATA, 2016. *Anais...*, Washington D.C., USA, 2016.

BARTLE, R. Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD Research*, v. 1, n. 1, p. 19, 1996. Disponível em: <<http://mud.co.uk/richard/hcds.htm>>. Acesso em: 28 jan. 2018.

BARTLE, R. Virtual Worlds: why people play. *Massively Multiplayer Game Development*, v. 2, p. 1-16, 2005. Disponível em: <<http://mud.co.uk/richard/VWWPP.pdf>>. Acesso em: 12 jan. 2018.

BARTLE, R. *Designing Virtual Worlds*. New Riders: Estados Unidos da América, 2003.

BARTLE, R. *What is MUD?* 1999. Disponível em: <<https://mud.co.uk/richard/wiremud.htm>>. Acesso em: 09 jul. 2017.

CAPLAN, S. E.; WILLIAMS, D.; YEE, N. Who plays, how much, and why? Debunking the stereotypical gamer profile. *Computers in Human Behavior*, v. 25, p. 1312-1319. 2008.

CNN FUTURE SUMMIT - VIRTUAL WORLDS. Direção e edição: CNN, 2007. Vídeo Online. Disponível em: <https://www.youtube.com/watch?v=BbY_zAOt-Ug>. Acesso em: 15 out. 2017.

DANKOFF, J. *Game Telemetry with DNA Tracking on Assassin's Creed*, 2014. Disponível em: <https://www.gamasutra.com/blogs/JonathanDankoff/20140320/213624/Game_Telemetry_with_DNA_Tracking_on_Assassins_Creed.php>. Acesso em 04 fev. 2018

DIXON, D. Player types and gamification. In: CHI WORKSHOP ON GAMIFICATION, 2011. Vancouver. *Anais...* 2011. Vancouver, Canada, 2011. p. 41-43.

DRACHEN A. What Are Game Metrics, 2012. Disponível em: <<https://gameanalytics.com/blog/what-are-game-metrics.html>>. Acesso em 13 nov. 2017

DRACHEN A. What Is Game Telemetry, 2012, disponível em: <<https://gameanalytics.com/blog/what-is-game-telemetry.html>>. Acesso em 13 nov. 2017.

DRACHEN, A. et al. Game data mining. In: EL-NASR, M. S.; DRACHEN, A.; CANOSSA, A. (Ed.). *Game Analytics: maximizing the value of player data*. Estados Unidos: Springer, 2013. p. 205–254.

DRACHEN, A. et al. Guns, Swords and Data: Clustering of Player Behavior in Computer Games in the Wild. In: IEEE CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND GAMES, *Anais...* 2012.

DRACHEN, A.; CANOSSA, A.; YANNANAKIS, G.N. Player modeling using self-organization in tomb raider: underworld. In: INTERNATIONAL SYMPOSIUM ON COMPUTATIONAL INTELLIGENCE AND GAMES, 7-10 Sept. 2009, *Anais ...* 2009.

DRACHEN, A.; EL-NASR, M. S.; CANOSSA, A. Game analytics – the basics. In: EL-NASR, M. S.; DRACHEN, A.; CANOSSA, A. (Ed.). *Game Analytics: Maximizing the Value of Player Data*. Estados Unidos: Springer, 2013. p. 13–40.

DUCHENEAUT, N., & NELSON, L., YEE, N. Online Gaming Motivations Scale: Development and Validation. In: Proceedings of CHI 2012, 2803-2806. 2012.

ESPINELI, M.; PLAGGE K. EVERY Game Release Date in 2017 disponível em: <<https://www.gamespot.com/articles/every-game-release-date-in-2017/1100-6446643/>>.

Acesso em 13 nov. 2017.

GOOGLE. Google Trends, 2017. Disponível em <<https://trends.google.com.br/trends/>> Acesso em: 07 jul. 2017.

HALOBI. Descriptive, Predictive, and Prescriptive Analytics Explained, 2017. Disponível em: <<https://halobi.com/blog/descriptive-predictive-and-prescriptive-analytics-explained/>>.

Acesso em 01 dez. 2017.

HAN, J.; KAMBER, M.; PEI, J. *Data Mining: concepts and techniques*. Estados Unidos: Morgan Kaufmann, 2012. p. 1-32

HERBERT, B., et al. An Investigation of Gamification Typologies for Enhancing Learner Motivation. In: INTERNATIONAL CONFERENCE ON INTERACTIVE TECHNOLOGIES AND GAMES. 16-17 Oct. 2014. *Anais...* 2014.

JORDAN, T. *Gamification 101: Richard Bartle player types*. 2014. Disponível em: <<https://repignite.com/2014/07/richard-bartle-player-types/>>. Acesso em: 20 abr. 2017.

KAGGLE. *World of Warcraft Avatar History Kaggle files*, 2015. Disponível em: <<https://www.kaggle.com/mylesoneill/warcraft-avatar-history>>. Acesso em: 27 jul. 2017.

KAGGLE. *World of Warcraft Avatar History Kaggle Kernels*. 2015. Disponível em: <<https://www.kaggle.com/mylesoneill/warcraft-avatar-history/kernels>>. Acesso em: 27 jul. 2017.

KOHONEN, T. *Self-Organizing Maps*. 3. ed. Estados Unidos: Springer, 2001.

LAZZARO, N. 4 Keys 2 Fun, 2013. Disponível em: <<http://www.nicolelazzaro.com/the4-keys-to-fun/>>. Acesso em 13 out 2011

Yeng-Ting Lee ; Kuan-Ta Chen ; Yun-Maw Cheng; Chin-Laung Lei. World of Warcraft Avatar History Dataset. In: *Proceedings of ACM Multimedia Systems 2011*, Santa Clara, CA, USA — February 23-25, 2011.

MACIEL, A. F. Uma interpretação nebulosa dos mapas de Kohonen. 2008. Dissertação (Mestrado em Modelagem Computacional)-Universidade Federal de Alagoas, Maceió, 2008.

MARCZEWSKI, A. The Intrinsic Motivation Ramp. 2013. Disponível em: <<https://www.gamified.uk/gamification-framework/the-intrinsic-motivation-ramp/>>. Acesso em 17 out 2017.

MARCZEWSKI, A. User Types. In: *Even Ninja Monkeys Like to Play: gamification, game thinking and motivational design*. Estados Unidos da América: CreateSpace Independent Publishing Platform, 2015. p. 65-80

MARCZEWSKI, A. *Using the Gamification User Types in the Real World*. 2013. Disponível em: <<https://www.gamified.uk/2013/12/09/using-gamification-user-types-real-world/>>. Acesso em: 17 out. 2017.

MCDONALD, E. The Global Games Market Will Reach \$108.9 Billion. In: *2017 With Mobile Taking 42%*. 2017. Disponível em: <<https://newzoo.com/insights/articles/the-global-games-market-will-reach-108-9-billion-in-2017-with-mobile-taking-42/>>. Acesso em: 20 jul. 2017

PLAYER TYPE THEORY: USES AND ABUSES | RICHARD BARTLE. Direção e produção: Casual Connect Europe, 2012. Online. Disponível em: <<https://www.youtube.com/watch?v=ZIZLbE-93nc>>. Acesso em: 10 mar. 2017.

REVISTABW Aprendizado de Máquina: aprendizado supervisionado. *Revista Brasileira de Web: Tecnologia*. Disponível em: <<https://www.revistabw.com.br/revistabw/aprendizagem-de-maquina-aprendizado-supervisionado/>>. Acesso em: 02 dez. 2017.

SCHELL, J. *The Art of Game Design: a book of lenses*. Estados Unidos: Morgan Kaufmann, 2008.

STATIA Number of available applications in the Google Play Store from December 2009 to September 2017 disponível em: <<https://www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play-store/>>. Acesso em: 13 nov. 2017.

STATIA. Value of the global video games market from 2011 to 2020 (in billion U.S. dollars). 2017. Disponível em: <<https://www.statista.com/statistics/246888/value-of-the-global-video-game-market/>>. Acesso em: 20 jul. 2017

SULLIVAN, J. Achievements and Gamerscore: Best Practices. 2009. Disponível em: <<http://www.voidcn.com/blog/baozi3026/article/p-3552500.html>>. Acesso em: 22 abr. 2017.

THE STATISTICS PORTAL. Number of World of Warcraft (WoW) subscribers from 1st quarter 2005 to 3rd quarter 2015 (in millions). Disponível em: <<https://www.statista.com/statistics/276601/number-of-world-of-warcraft-subscribers-by-quarter/>> Acesso em: 13 nov. 2017.

WORTH, N. C.; BOOK, A. S. Personality and behavior in a massively multiplayer online role-playing game. *Computers in Human Behavior*, 2014. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0747563214003446>>. Acesso em: 12 dez. 2017.

WOWWIKI. *Timeline* (World of Warcraft). Disponível em: <[http://wowwiki.wikia.com/wiki/Timeline_\(World_of_Warcraft\)](http://wowwiki.wikia.com/wiki/Timeline_(World_of_Warcraft))>. Acesso em: 08 mar. 2018.

YANNAKAKIS, G. N.; TOGELIUS, J. A Panorama of Artificial and Computational Intelligence in Games, *IEEE Transactions on Computational Intelligence and AI in Games*, v. 7, n. 4, dez. 2015

YEE et al. *Introverted Elves and Conscientious Gnomes: the expression of personality in World of Warcraft*. Proceedings of CHI 2011, p. 753-762. 2011

YEE, N. The demographics, motivations and derived experiences of users of Massively-Multiuser Online Graphical Environments. *Presence: Teleoperators and Virtual Environments*, 15. 2006.

YEE, N. Motivations for Play in Online Games. *CyberPsychology and Behavior*, v. 9, p. 772-775, 2006.

ZENN, J. *Maximizing the value of Player Data*, 2012. Disponível em: <<https://gameanalytics.com/blog/maximizing-the-value-of-player-data.html>>. Acesso em: 10 nov. 2017.