

Explorando Estratégias para Estimativa da Relevância de Opiniões no Domínio de Jogos Eletrônicos

Exploring Strategies for Estimating Opinion Relevance in the Electronic Games Domain

Ellaine Barros*¹, Allan Diego Silva Lima¹, Enok Tavares¹, Allyson Silva¹

*ellaine.dayane@gmail.com

¹Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco

RESUMO

Este artigo apresenta um estudo sobre aspectos que tornam uma opinião sobre um jogo eletrônico relevante para usuários de redes sociais na internet. Para tal, seu conteúdo baseou-se no seguinte questionamento: dada a grande quantidade de opiniões presentes na internet, como encontrar aquelas que são mais relevantes para um determinado usuário? Tentando responder a esta pergunta, este artigo relata uma série de experimentos sobre o conceito de similaridade entre usuários de redes sociais e a polaridade dos textos das opiniões para a estimativa da relevância das mesmas. Neste contexto, foram exploradas diversas variações dos conceitos, a saber: Similaridade de Idade, a Similaridade de Interesses e Similaridade de Amigos, e, no âmbito da polaridade de opiniões foram estudadas a positividade e a negatividade das mesmas. A efetividade de cada uma das estratégias foi avaliada através de um corpus de opiniões sobre jogos eletrônicos. Durante os experimentos, no âmbito da similaridade, o modelo que se destacou foi o *SubjectSimilarity*, que trata opiniões de pessoas com interesses em comum dentro de uma rede social como mais relevantes. Já no âmbito da polaridade, o modelo que atingiu melhor resultado foi o *SentiNetPositiveOrientation*, que trata opiniões positivas como as mais relevantes.

Palavras-Chaves: Mineração de Opiniões, Recuperação de Informação, Jogos Eletrônicos, Ranqueamento de Opiniões.

ABSTRACT

This paper presents a study about witch aspects make an opinion about an electronic game relevant to social network users on the internet. Therefore, its content is based on the following research question: given de immense amount of opinions on the internet, how to find out those that are more relevant to a specific reader? To answer this question, thispaper

describes a series of experiments about the concept of similarity between social network users and the polarity of the opinions' texts as heuristics to estimate de relevanceof an opinion. In this context, distinct variations the following concepts were explored: similarity of age; similarity of interests; similarity of friends; and, on the polarization domain, opinions' positivity and negativity. The effectiveness of each strategies was evaluated by using an opinion corpus about electronic games. During the experiments, when it comes to similarity, the best strategy proposed is called SubjectSimilarity and it prioritizes opinions from authors with more interests in common. Furthermore, when it comes to the polarity, the best proposed strategy is called SentiNetPositiveOrientation and it prioritizes positive opinions as the more relevant ones.

Keywords: Opinion Mining, Information Retrieval, Electronic Games, Opinion Ranking.

1. Introdução

Devido a crescente importância e presença das redes sociais digitais na sociedade contemporânea, buscas por informações presentes nestas têm sido cada vez mais populares. Neste contexto, um dos problemas presentes em um ramo da ciência da computação conhecido como Mineração de Opiniões (Liu, 2012; Lima; Sichman, 2015) trata de como identificar as opiniões presentes na internet que são mais relevantes para um determinado leitor.

Neste âmbito, o presente artigo traz propostas e testes de heurísticas capazes de estimar o quão relevante uma opinião é, visando desta forma, visando contribuir para a construção de soluções para o problema descrito no parágrafo anterior.

Atualmente, o problema da relevância de opiniões torna-se importante a medida quem sua solução possui várias aplicações diretas em serviço que hoje são oferecidos na internet. Possivelmente, a sua aplicação mais direta seria na construção de um engenho de busca de opiniões. Além disto, também seria possível aplicar o modelo em redes sociais tradicionais como o Twitter e o Facebook uma vez que parte dos conteúdos da mesma é formado por opiniões sobre os mais diversos assuntos.

Entretanto, antes que qualquer aplicação direta de heurísticas para estimar a relevância de opiniões é preciso defini-las formalmente e testá-las com conjuntos de opiniões advindas de redes sociais. Desta forma, os resultados aqui apresentados foram realizados fazendo o uso do *Social Opinion Relevance Model* (SORM) (Lima; Sichman, 2015) um projeto que propôs um conjunto de parâmetros (ou heurísticas) capazes de estimar a relevância de opiniões para um dado membro de uma rede social.

Assim, o presente artigo realiza uma extensa reavaliação de parte dos parâmetros originais do SORM procurando por formas alternativas e nunca antes testadas de computar aspectos da relevância de uma opinião que atualmente estão presentes no modelo, em busca de resultados superiores do que os obtidos originalmente no modelo para compor a sua segunda versão.

Dentre os parâmetros do SORM, aqueles relacionados ao conceito de similaridade, foram selecionados como ponto de partida para os estudos realizados, buscando assim explorar melhores formas de se estimar a Similaridade de Idade, a Similaridade de Interesses entre membros de redes sociais, além da proposição de um novo parâmetro ao SORM: a Similaridade de Amigos. Os resultados deste projeto foram obtidos através de experimentos utilizando opiniões sobre Jogos Eletrônicos, tais resultados têm agregado valor aos experimentos desenvolvidos primariamente com o SORM, e desta forma, ajudado a entender e explicar devidamente as lacunas que ele deixou em sua primeira etapa.

Ao longo desse artigo, serão abordados na próxima seção os trabalhos relacionados; em sua terceira seção, a metodologia do projeto; na seção quatro serão apresentados e discutidos os resultados e os experimentos; por fim, na seção cinco serão retratadas as conclusões e as possibilidades de trabalhos futuros.

2. Trabalhos Relacionados

É possível observar que o conceito de relevância de opinião é comumente aplicado com o objetivo de estimar a relevância de documentos (Mishne, 2006; Zhang; Ye, 2008; Huang; Croft, 2009). Ou seja, projetos nesta área utilizam características de opiniões presentes em documentos para estimar a relevância dos mesmos. Assim, nestes projetos, opiniões são tratadas como membros de segunda classe, uma vez que não se tem como objetivos medir suas relevâncias, mas sim as dos documentos onde as mesmas foram escritas.

Mais recentemente, Projetos na área (Orimaye, *et al.*, 2012; Xu, *et al.*, 2012) chegam até a explorar aspectos linguísticos das opiniões, ou seja, passaram a avaliar o conteúdo dos seus textos. Porém, ainda visando estimar a relevância de documentos, sem, desta forma, possuir como objetivo principal a relevância de opiniões.

Neste contexto, foi desenvolvido o *Social Opinion Relevance Model* (SORM) (Lima; Sichman, 2013) como resultado de um projeto de doutorado na Universidade de São Paulo (USP). O principal objetivo do SORM é estudar como realizar a estimativa personalizada de quais são as opiniões mais relevantes para um determinado usuário de uma rede social.

O SORM é composto por um conjunto de parâmetros que são utilizados para estimar a relevância de uma opinião. Por exemplo, um dos parâmetros que compõe o SORM baseia-se na similaridade entre um consumidor (leitor) e o autor de uma opinião sendo representado por uma estimativa entre a quantidade de interesses em comum que ambos compartilham. Durante a concepção do SORM devido a restrições de tempo e recursos, não foi possível realizar uma detalhada e profunda análise sobre as melhores estratégias para se computar cada um dos parâmetros do modelo. E é com o objetivo de preencher justamente esta lacuna que o presente artigo foi escrito.

3. *Materiais e Métodos*

A metodologia do trabalho de pesquisa que foi aplicada no projeto está dividida em duas etapas, a saber: (1) desenvolvimento de novas fórmulas para os parâmetros do SORM; (2) avaliação das fórmulas geradas na etapa anterior através do uso de métricas de recuperação de informação.

O desenvolvimento de novas fórmulas para os parâmetros do SORM (Figura 1) surgiu através de duas atividades complementares. A primeira foi um processo de *Brainstorming* onde os pesquisadores foram incentivados a usar sua criatividade e conhecimento empírico para encontrar novas formas de computar os parâmetros do modelo. Durante este processo foi estudada a seguinte questão: “O que faz uma opinião relevante para uma determinada pessoa?”. Os resultados de ambas as atividades passaram por um processo de estudo e seleção chamado aqui de “Avaliação de Viabilidade”, onde foi verificado se as estratégias propostas eram passíveis de serem computadas através dos dados comumente presentes em redes sociais.

Figura 1. Gerando variações dos parâmetros presentes no SORM.

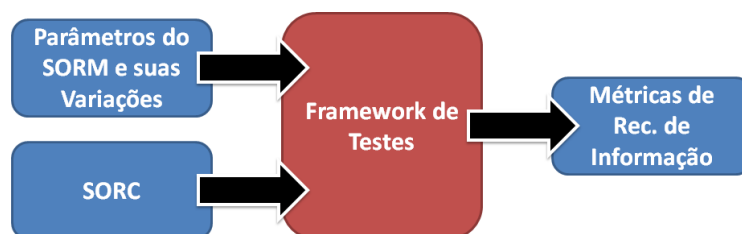


Fonte: Elaborada pelos autores.

A avaliação de fórmulas para relevância de opinião, Figura 2, baseia-se na utilização do *Framework* de Testes do SORM para avaliação dos parâmetros que estimam a relevância de uma opinião. Ele faz uso de métricas de tradicionais da área de

Recuperação de Informação (Manning *et al.*, 2008) visando identificar qual dos parâmetros que capaz de estimar melhor a relevância de uma opinião.

Figura 2. Comparando os novos e os antigos parâmetros do modelo.



Fonte: Elaborada pelos autores.

Já o corpus utilizado para testar os parâmetros foi o *Social Opinion Relevance Corpus* (SORC) Games-01 (Lima; Sichman, 2013), um corpus de opiniões sobre jogos eletrônicos. Este corpus é composto por um conjunto de 50 casos de teste, onde cada caso de testes é composto por: uma coleção de opiniões sobre um jogo específico; um usuário de uma rede social; e um subconjunto das opiniões sobre o jogo que foram julgadas pelo usuário.

Dentro do *framework* de testes cada parâmetro é testado em todos os 50 casos de teste. Para cada caso teste, o parâmetro é utilizado como função de ordenação das opiniões de acordo com sua relevância estimada, simulando assim como um engenho de busca tradicional funciona ao receber uma consulta (porém, neste caso, a consulta não visa encontrar documentos, mas sim opiniões).

Os parâmetros capazes colocar as opiniões avaliadas como relevantes pelo usuário nos primeiros lugares da ordenação são considerados os melhores. Desta forma, as métricas utilizadas na avaliação, de forma geral, exibem um melhor resultado quando o modelo é capaz de priorizar as opiniões marcadas como relevante pelo usuário associado ao caso de testes.

As diversas métricas utilizadas para avaliação dos parâmetros foram selecionadas através de uma intensiva revisão bibliográfica realizada ainda durante o desenvolvimento da primeira versão do SORM (Lima; Sichman, 2015). Dentre elas é possível destacar a *Average Precision* (AP) que computa a média da precisão do parâmetro em cada um dos casos de testes, além das métricas *QMeasure* (QMea.) e *Normalized Discounted Cumulative Gain* (NDCG) que seguem a mesma linha, porém não consideram a relevância

como um conceito binário (relevante, não relevante), mas sim um conceito com valores intermediários de relevância.

Como no SORC Games-01 as opiniões foram classificadas pelos usuários entre muito relevantes, relevantes, pouco relevantes e irrelevantes, a principal métrica de avaliação utilizada para definir qual o melhor parâmetro para estimar a relevância de opiniões é a média harmônica entre métricas *QMeasure* e *Normalized Discounted Cumulative Gain*. Assim, é possível identificar como o melhor parâmetro aquele capaz de prover bons resultados em ambas as métricas.

4. Resultados e Discussão

Durante a primeira etapa da pesquisa relatada neste artigo, diversas foram as discussões sobre como melhorar a primeira versão do SORM, dentre elas podemos destacar: (1) As diretrizes que nortearam a criação de novas fórmulas para os parâmetros do SORM; (2) A definição das novas fórmulas candidatas a comporem a próxima versão do SORM; (3) Execução de um conjunto de testes para verificar quais fórmulas se adequam melhor ao domínio de Jogos Eletrônicos.

O processo de pesquisa para melhoria dos parâmetros do SORM iniciou-se com a definição três características que seriam importantes para as fórmulas do modelo:

- 1. Quanto mais próximas forem as características, maior será o valor dado pela fórmula.** Por exemplo, no contexto da similaridade de idade de membros de redes sociais, suponha que o autor de uma opinião possua 25 anos de idade e que suas opiniões terão a relevância estimada para uma pessoa de X anos de idade, quanto mais próximo X for de 25, maior deverá ser o valor dado pelo parâmetro do modelo para relevância da opinião;
- 2. O resultado da relevância dado pelo parâmetro deve estar no intervalo numérico real inclusivo entre 0 e 1.0.** Isto se deve ao fato de que ao distribuímos o valor entre este intervalo, caracterizamos o parâmetro como uma probabilidade, assim o seu resultado significa o quão provável é uma opinião seja relevante para um determinado usuário;
- 3. A inversão dos papéis entre usuário e consumidor da opinião deve gerar o mesmo resultado (i.e. a mesma estimativa de relevância).** Como a maioria dos parâmetros do SORM não atende a esta propriedade, a inclusão de parâmetros que a respeita torna possível o estudo da hipótese de que a relevância de uma opinião não é associativa. Ou seja, a relevância de uma

opinião do autor *A* para o consumidor *C* não é a mesma de uma opinião que tenha *C* como autor e *A* como consumidor.

Visando testar as fórmulas que foram criadas com base nas diretrizes acima, utilizamos o SORC Games01 do *Social Opinion Relevance Corpus Project* (Lima; Sichman, 2013) um corpus com opiniões sobre Jogos Eletrônicos que foi criado na Universidade de São Paulo durante o projeto de pesquisa se gerou o SORM. O SORC Games01 está disponível na internet¹, de forma que qualquer pesquisador interessado possa baixá-lo e utilizá-lo gratuitamente.

Durante o projeto dos experimentos, foram definidos os resultados obtidos com as novas fórmulas para a estimativa da relevância de opinião seriam estudados e comparados com resultados de dois paradigmas distintos, a saber: (1) comparação com modelos relativamente simples, porém aplicados tradicionalmente em sistemas de recuperação de informação; (2) comparação de referência com fórmulas originalmente presentes no SORM.

Os dois modelos comumente presentes em sistemas de recuperação de informação selecionados para o experimento foram o *RareWordsFrequency (RWF)* e o *LinearLongevity (LL)*. O modelo *RareWordsFrequency* baseia-se na raridade das palavras, assim quanto mais palavras raras uma opinião possui maior será a sua relevância, este modelo inspira-se em uma estratégia consolidada na recuperação de informação e aplicada na composição do *TF-IDF (Term Frequency – Inverse Document Frequency)* (Manning et al., 2008). Já o modelo *LinearLongevity* prioriza as opiniões mais recentes, ou seja, quanto mais nova for a data da publicação da opinião, maior será a sua relevância.

Os modelos são comparados através de um experimento onde as opiniões presentes no SORC são utilizadas como entrada para que os modelos possam estimar suas relevâncias. Intuitivamente, quanto mais opiniões relevantes um modelo for capaz de identificar, melhor o seu desempenho. O desempenho dos modelos é medido através da média de 50 casos de testes prestes no SORC Games01. Cada caso de teste é composto por um conjunto de opiniões sobre um determinado jogo eletrônico, onde um subconjunto destas opiniões teve cada uma das suas opiniões classificada entre não relevante; pouco relevante; relevante; e muito relevante por um consumidor (usuário de uma rede social). Ao final do experimento são computadas métricas de avaliação de sistemas de recuperação de informação (Manning et al., 2008; Lima; Sichman, 2013).

¹ <https://allanlima.wordpress.com/category/sorc/>

4.1. Similaridade de Idade

A estimativa da similaridade de idade se baseia na diferença de idade entre um autor e um consumidor de uma opinião. Intuitivamente, quanto menor a diferença entre as idades, mais relevante é a opinião para o consumidor. Por exemplo, esperamos que as pessoas da mesma geração (em torno da mesma idade) pensem da mesma forma.

Dentro do conjunto de parâmetros do SORM, a estimativa da similaridade entre as idades de usuários em redes sociais, era a estratégia que se baseava em uma operação de subtração entre as idades, originalmente chamada de *AgeSimilarity*. Neste artigo propomos mais duas novas estratégias: a *AgeSimilarityByDivision* e a *AgeSimilarityBySwap*. Tanto a estratégia original, como as duas novas propostas serão descritos a seguir. Assim, dado o autor da opinião *A* e consumidor da opinião *U*; e dadas as idades do autor da opinião (*AgeA*) e do consumidor da opinião (*AgeC*):

AgeSimilarity (AS): A fórmula proposta na primeira versão do SORM (1) que utiliza como parâmetros as idades e tem como seu resultado uma fórmula baseada na subtração das duas idades. Quanto mais próximas são as idades, a diferença entre elas será mais próxima de zero, portanto, o valor do resultado será maior, de acordo com a fórmula abaixo. A constante *k* não foi definida propositalmente visando permitir a regulação do crescimento do valor do parâmetro de acordo com as necessidades dos experimentos.

$$AgeSimilarity(A,U) = \max \left(0, 1 - \left(\frac{|AgeA - AgeC|}{k} \right) \right) \quad (1)$$

AgeSimilarityByDivision (ASBD): Fórmula que se baseia na divisão da soma das idades pelo módulo da diferença, assim como ilustrado na Equação 2. Note que quando as idades são idênticas, o resultado da função é 1.

$$AgeSimilarityByDivision(A,U) = \frac{1}{1 + |AgeA - AgeC|} \quad (2)$$

AgeSimilarityBySwap (ASBS): Baseia-se na divisão da menor idade pela maior idade (Equação 3) note a presença do +1 no denominador da divisão existe para prevenir uma divisão por 0, uma vez que em redes sociais nem sempre os usuários informam todos os dados em seus perfis.

$$AgeSimilarityBySwap(A,U) = \frac{\min(AgeA, AgeC)}{\max(AgeA, AgeC) + 1} \quad (3)$$

Inicialmente, ao estudarmos os resultados obtidos (Tabela 1), buscamos identificar qual dos modelos obteve os melhores resultados nas métricas *Q-Measure'* e no *NDCG'*, que são as principais métricas selecionadas para avaliação durante os experimentos do SORM (Lima; Sichman, 2014). Porém, como melhoria sobre o método de avaliação dos resultados utilizados no SORM, definimos a *F-Metric* (a média harmônica *Q-Measure'* e *NDCG'*) como nossa principal métrica de avaliação.

Tabela 1. Resultados dos experimentos sobre a Similaridade de Idade, quanto maior melhor o desempenho do modelo na métrica.

Modelo	NDCG	NDCG'	R-Prec.	R-Prec.'	Bpref	Bpref'	Q-Mea.	Q-Mea.'	AP	AP'	F-Metric
RWF	0.0297	0.2442	0.0409	0.2700	0.4191	0.2700	0.0093	0.0340	0.0101	0.2650	0.0597
LL	0.0397	0.2342	0.0490	0.2167	0.4269	0.2311	0.0205	0.0433	0.0214	0.2133	0.0731
AS	0.0397	0.2342	0.0490	0.2167	0.4269	0.2311	0.0205	0.0433	0.0214	0.2133	0.0731
ASBD	0.0281	0.2049	0.0536	0.2133	0.4239	0.2206	0.0197	0.0444	0.0203	0.2011	0.0730
ASBS	0.0276	0.2215	0.0508	0.2133	0.4224	0.2206	0.0185	0.0429	0.0190	0.2011	0.0719

Fonte: Elaborada pelos autores.

Sob tal ótica, *AgeSimilarityBySwap* e *AgeSimilarityByDivision* se destacaram nos resultados das métricas *Q-Measure'* e *NDCG'* respectivamente, o que representa evidências de que ambas podem ser alternativas superiores para estimar a relevância de opinião quando comparadas ao modelo de similaridade de idade originalmente presente no SORM (*AgeSimilarity*). Porém, ao analisarmos todas as métricas da tabela, é possível verificar que os modelos que demonstraram melhores valores no número de métricas são a *LinearLongevity* e a *AgeSimilarity*, sendo os melhores modelos em 5 das 11 métricas. Por conta desta dificuldade em superar os resultados de estratégias comumente aplicadas em sistemas de recuperação, os resultados, apesar de promissores, ainda não podem ser categorizados como conclusivos. Por tanto, no contexto da similaridade de idade, ainda há importantes estudos a serem desenvolvidos buscando evidências para uma resposta mais incisiva sobre qual a melhor fórmula para computar a relevância com base a similaridade de idade.

4.2. Similaridade de Interesses

No SORM, a Similaridade entre autor e consumidor é representada por uma estimativa entre a quantidade de interesses em comum que um consumidor compartilha com um autor de uma opinião. Quanto mais interesses eles têm em comum, mais relevantes às opiniões do autor são para um consumidor. Por exemplo, no domínio de livros, quando mais livros dois usuários de uma rede social leram em comum, maior seria a relevância da opinião de um para o outro.

Assim, dada a função $Subject(U)$, que retorna o conjunto de interesses de um usuário U de uma rede social e sejam C e A dois usuários de uma rede social, onde C exerce o papel de consumidor da opinião e A o papel do autor da opinião:

SubjectSimilarity (SB): calcula cardinalidade da interseção dos interesses desses dois tipos de usuários e divide essa quantidade pelos interesses dos consumidores, assim como ilustrado na equação 4.

$$SubjectSimilarity(C,A) = \frac{|Subjects(C) \cap Subjects(A)|}{|Subjects(C)|} \quad (4)$$

AssociativeSubjectSimilarity (ASS) Fórmula que calcula a intersecção dos interesses de um consumidor com os interesses de um autor, divide pela cardinalidade do interesse de ambos, assim como ilustrado na equação 5. A soma na equação pretende fazer com que ela atenda a propriedade da associatividade.

$$AssociativeSubjectSimilarity(C,A) = \frac{|Subjects(C) \cap Subjects(A)|}{|Subjects(C)| \times |Subjects(A)|} \quad (5)$$

Novamente, ao analisarmos os resultados obtidos na (Tabela 2), os resultados indicam que $SubjectSimilarity$ e $SubjectAssociativeSimilarity$ obtiveram melhores resultados nas métricas $Q-Measure'$ e $NDCG'$ respectivamente. Entretanto, ao analisarmos todos os parâmetros da tabela, percebemos que o modelo que demonstrou melhor valor no número de métricas é a $Similarity$, sendo o melhor modelo em nove das onze métricas. Esta diferença é um importante indicativo de que, no contexto de jogos eletrônicos, a associatividade não é uma propriedade necessária na estimativa da relevância de opiniões com base em interesses em comum.

Tabela 2. Resultados dos experimentos sobre a Similaridade de Interesses, quanto maior melhor o desempenho do modelo na métrica.

Model	NDCG	NDC G'	R-Prec.	R-Prec.'	Bpref	Bpref'	Q-Mea.	Q-Mea.'	AP	AP'	F-Metric
RWF	0.0297	0.2442	0.0409	0.2700	0.4191	0.2700	0.0093	0.0340	0.0101	0.265	0.0597
LL	0.0397	0.2342	0.049	0.2167	0.4269	0.2311	0.0205	0.0433	0.0214	0.2133	0.0731
ASS	0.0417	0.3049	0.0620	0.2933	0.4311	0.3056	0.0223	0.0549	0.0229	0.2861	0.0992
SS	0.0472	0.3018	0.0680	0.3000	0.4360	0.3167	0.0249	0.0593	0.0253	0.2894	0.0930

Fonte: Elaborada pelos autores.

Por fim, dados os resultados, ainda não é possível afinar que há evidências fortes que validem ou não a hipótese de que associatividade não é importante no contexto da similaridade de interesses. Ficando esta análise como uma lacuna aberta na ciência para trabalhos futuros.

4.3. Similaridade de Amigos

A similaridade de amigos é baseada na intersecção entre os conjuntos de amigos de outros dois usuários de uma rede Social. Quanto mais amigos em comum, mais relevantes são as opiniões do autor para um dado consumidor.

Formalmente, seja *Friends (U)* é uma função que dado um usuário *U* de uma rede social, resulta no conjunto de amigos que o usuário possui na rede, sejam *C* e *A* dois usuários de uma rede social, onde *C* exerce o papel de consumidor da opinião e *A* o papel do autor da opinião:

FriendSimilarity (FS): Calcula a cardinalidade da intersecção entre os amigos do autor e do consumidor, este resultado é então dividido pelo tamanho do conjunto de amigos do consumidor, assim como ilustrado na equação 6.

$$FriendSimilarity(C,A) = \frac{|Friends(C) \cap Friends(A)|}{|Friends(C)|} \quad (6)$$

AssociativeFriendSimilarity (AFS): É um parâmetro que calcula a intersecção dos amigos de um consumidor com os amigos de um autor, divide pelos amigos do consumidor

e depois pelos amigos do autor, então soma esses resultados e faz uma média, assim como ilustrado na equação 7.

$$FriendAssociativeSimilarity(C,A) = \frac{|Friends(C) \cap Friends(A)|}{Friends(C) \times Friends(A)} \quad (7)$$

Os resultados encontrados (Tabela 3) mostram que *FriendSimilarity* e *FriendAssociativeSimilarity* obtiveram melhores resultados nas métricas *Q-Measure'* e *NDCG'* respectivamente. Entretanto, ao analisarmos todos os parâmetros da tabela, percebemos que o modelo que demonstrou melhor valor no número de métricas é a *FriendSimilarity*, sendo o melhor modelo em cinco das onze métricas. Além disso, *FriendSimilarity* demonstrou o melhor resultado na *F-Metric* o que, segundo os critérios definidos para os experimentos significa que foi possível encontrar evidências de que ela é melhor que as demais na estimativa da relevância de opiniões sobre jogos eletrônicos.

Tabela 3. Resultados dos experimentos sobre a Similaridade de Amigos, quanto maior melhor o desempenho do modelo na métrica.

Model	NDCG	NDCG'	R-Prec.	R-Prec. '	Bpref	Bpref'	Q-Mea.	Q-Mea.'	AP	AP'	F-Metric
RWF	0.0297	0.2442	0.0409	0.2700	0.4191	0.2700	0.0093	0.0340	0.0101	0.2650	0.0597
LL	0.0397	0.2342	0.0490	0.2167	0.4269	0.2311	0.0205	0.0433	0.0214	0.2133	0.0731
FS	0.0368	0.2432	0.0575	0.2533	0.4284	0.2606	0.0217	0.0497	0.0221	0.2461	0.0826
FAS	0.0291	0.2451	0.0533	0.2600	0.4233	0.2667	0.0184	0.0448	0.0185	0.2444	0.0757

Fonte: Elaborada pelos autores.

Por fim, assim como na similaridade de interesses, dados os resultados, ainda não é possível afinar que há evidências fortes que validem ou não a hipótese de que associatividade não é importante no contexto da similaridade de amigos.

4.4. Polaridade de Opiniões

Os modelos de polaridade de opiniões tentam encontrara uma relação entre a polaridade de uma opinião (sua positividade ou negatividade) e a sua relevância. Assim, foram definidas estratégias distintas de forma a explorar diversas maneiras de se relacionar a positividade e/ou negatividade de uma opinião com a relevância da mesma para um usuário de uma rede social.

Durante a formalização das estratégias, vamos utilizar O como o conjunto das palavras de uma opinião; $Positivity(w)$ uma função que dada uma palavra w tem como resultado a positividade de w ; e $Negativity(w)$ uma função que dada uma palavra w tem como resultado a negatividade de w . Para os experimentos com estes modelos para a implementação das funções $Positivity$ e $Negativity$ foi utilizado o dicionário de polaridade de palavras *SentiWordNet* (Baccianella, et al., 2010).

SentNetPositiveOrientation (verifica a positividade das palavras) e *SentNetNegativeOrientation* (verifica a negatividade das palavras) foram utilizadas como referência para os resultados dos experimentos. Durante os experimentos as seguintes estratégias foram testadas:

SentiNetPositiveOrientation (SNPO): Prioriza opiniões positivas, ou seja, quando mais palavras com orientação positiva uma opinião possui, mais relevante ela é considerada.

$$SentNetPositiveOrientation(O) = \frac{\sum_{w \in O} Positivity(w)}{|O|} \quad (8)$$

SentiNetNegativeOrientation (SNNO): Prioriza opiniões negativas, ou seja, quando mais palavras com orientação negativa uma opinião possui, mais relevante ela é considerada.

$$SentNetNegativeOrientation(O) = \frac{\sum_{w \in O} Negativity(w)}{|O|} \quad (9)$$

SentiNetOrientation (SNO): Prioriza opiniões com maior diferença entre a positividade e a negatividade. Ou seja, esta estratégia prioriza opiniões que são majoritariamente positivas ou negativas.

$$SentiNetOrientation(O) = \frac{\sum_{w \in O} |Positivity(w) - Negativity(w)|}{|O|} \quad (10)$$

SentNetPositiveOrientationByDivision (SNPOBD): Prioriza opiniões que contenham mais palavras com orientação positivas do que negativas.

$$\begin{aligned} & \text{SentNetPositiveOrientationByDivision}(O) \\ &= \frac{\text{SentNetPositiveOrientation}(O)}{\text{SentNetPositiveOrientation}(O) + \text{SentNetNegativeOrientation}(O) + 1} \end{aligned} \quad (11)$$

SentNetNegativeOrientationByDivision (SNNODB): Prioriza opiniões que contenham mais palavras com orientações negativas do que positivas.

$$\begin{aligned} & \text{SentNetNegativeOrientationByDivision}(O) \\ &= \frac{\text{Negativity}}{\text{SentNetPositiveOrientation}(O) + \text{SentNetNegativeOrientation}(O) + 1} \end{aligned} \quad (11)$$

SentNetOrientationByDivision (SNOBD): prioriza opiniões que tenham um balanço sobre as polaridades, ou seja, opiniões que seriam consideradas como neutras.

$$\begin{aligned} & \text{SentNetOrientationByDivision}(O) \\ &= \frac{1}{(|\text{SentNetPositiveOrientation}(O) - \text{SentNetNegativeOrientation}(O)| + 1)} \end{aligned} \quad (12)$$

SentNetOrientationBySwap (SNOBS): dividi o menor valor entre a positividade e negatividade da opinião pelo maior valor, note que esta estratégia também prioriza opiniões podem se consideradas como neutras.

$$\begin{aligned} & \text{SentNetOrientationBySwap}(O) \\ &= \frac{\min(\text{SentNetPositiveOrientation}(O), \text{SentNetNegativeOrientation}(O))}{\max(\text{SentNetPositiveOrientation}(O) + 1, \text{SentNetNegativeOrientation}(O) + 1)} \end{aligned} \quad (13)$$

Inicialmente, ao analisarmos os resultados obtidos na (Tabela 4) procuramos identificar qual dos modelos obteve os melhores resultados nas métricas *Q-Measure'* e *NDCG'*, que são as principais métricas selecionadas para avaliação durante os experimentos do SORM (Lima; Sichman, 2013). No entanto, *SentiNetPositiveOrientation* e *SentNetNegativeOrientationByDivision* obtiveram melhores resultados nas métricas *Q-Measure'* e *NDCG'* respectivamente. Entretanto, ao analisarmos todos os parâmetros da tabela, percebemos que o modelo que demonstrou melhor valor no número de métricas é a

SentiNetPositiveOrientation, sendo o melhor modelo em sete das onze métricas além de se destacar na principal métrica de avaliação (*F-Metric*).

Tabela 4. Resultados dos experimentos sobre a Polaridade de opiniões, quanto maior melhor o desempenho do modelo na métrica.

Model	NDC _G	NDCG _'	R-Prec.	R-Prec. _'	Bpref	Bpref _'	Q-Mea.	Q-Mea _'	AP	AP _'	F-Metric
RWF	0.0297	0.2442	0.0409	0.2700	0.4191	0.2700	0.0093	0.0340	0.0101	0.2650	0.0597
LL	0.0397	0.2342	0.0490	0.2167	0.4269	0.2311	0.0205	0.0433	0.0214	0.2133	0.0731
SNO	0.0641	0.3155	0.0644	0.3700	0.4401	0.3750	0.0269	0.0558	0.0300	0.3700	0.0948
SNPO	0.0926	0.2777	0.0700	0.3000	0.4488	0.3000	0.0405	0.0607	0.0460	0.3000	0.0997
SNNO	0.0427	0.2961	0.0528	0.3400	0.4281	0.3463	0.0163	0.0439	0.0181	0.3350	0.0765
SNNODB	0.0409	0.3170	0.0535	0.3533	0.4287	0.3628	0.0160	0.0448	0.0177	0.3483	0.0786
SNNODB	0.0430	0.2528	0.0483	0.2933	0.4266	0.2956	0.0142	0.0397	0.0160	0.2911	0.0687
SNPOBD	0.0888	0.2772	0.0678	0.3000	0.4481	0.3000	0.0393	0.0586	0.0445	0.3000	0.0968
SNOBS	0.0441	0.2642	0.0447	0.2733	0.4276	0.2733	0.0158	0.0376	0.0171	0.2678	0.0659

Fonte: Elaborada pelos autores.

5. Conclusões

Durante este artigo foram apresentadas e avaliadas diversas estratégias para estimativa da relevância de opiniões que serão futuramente integradas ao SORM. Ao executarmos os experimentos com os novos parâmetros obtivemos indícios de que parte das novas fórmulas propostas para integrar o SORM 2.0 tem melhor desempenho nas métricas de avaliação em Recuperação de Informação utilizadas nos experimentos. Entretanto, mais experimentos ainda se fazem necessários.

Durante todo os experimentos realizados para este artigo foram avaliadas diversas métricas de recuperação de informação e criada a *F-Metric* como principal forma de definir qual o melhor modelo. Entretanto, os resultados dos experimentos apresentados neste artigo poderiam ser melhorados a através da aplicação de um teste de hipótese, uma vez que o valor de cada métrica é a média de 50 casos de teste. Desta forma, ainda é preciso definir que tipo de teste de hipótese deve ser aplicado especificamente para este tipo de teste.

Por fim, como trabalhos futuros, todos os resultados apresentados neste artigo serão integrados ao SORM. Assim, as melhores estratégias identificadas neste artigo irão substituir suas estratégias similares presentes no SORM. Além disto, as novas estratégias serão adicionadas ao SORM. Ao final deste processo será possível utilizar algoritmos de aprendizado de máquina combinar as melhores estratégias aqui descritas com as demais estratégias presentes no SORM gerando um modelo único.

Referências

GILAD MISHNE. Multiple Ranking Strategies for Opinion Retrieval in Blogs - The University of Amsterdam at the 2006 TREC Blog Track. **Text REtrieval Conference**, 1 jan. 2006.

HUANG, X.; CROFT, W. B. A Unified Relevance Model for Opinion Retrieval. **Proceeding of the 18th ACM Conference on Information and Knowledge Management - CIKM '09**, v. 1, p. 947–956, 2009.

LIMA, A.; SICHMAN, J. **S.O.R.M.: Social Opinion Relevance Model**. Thesis—Universidade de São Paulo: [s.n.], 2015.

LIMA, A.; SICHMAN, J. S. SORM: A Social Opinion Relevance Model. **2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)**, 1 ago. 2014.

LIU, B. Sentiment Analysis and Opinion Mining. **Synthesis Lectures on Human Language Technologies**, v. 5, n. 1, p. 1–167, 23 maio 2012.

MANNING, C.; PRABHAKAR, R.; SCHÜTZE, H. **Introduction to Information Retrieval**. 1. ed. [s.l.] Cambridge University Press, 2008. v. 1

ORIMAYE, S. O.; ALHASHMI, S. M.; SIEW, E.-G. Can predicate-argument structures be used for contextual opinion retrieval from blogs? **World Wide Web**, v. 16, n. 5-6, p. 763–791, 31 maio 2012.

STEFANO BACCIANELLA; ESULI, A.; SEBASTIANI, F. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. **Language Resources and Evaluation**, 1 maio 2010.

XU, X.; TAN, S.; LIU, Y.; CHENG, X.; LIN, Z.; GUO, J. Find me Opinion Sources in Blogosphere: a Unified Framework for Opinionated Blog Feed Retrieval. **Proceedings of the Fifth ACM International Conference on Web Search and Data Mining**, pp. 583-592. ACM, 2012.

Zhang, M.; & Ye, X.. A Generation Model to Unify Topic Relevance. **31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval**, pp. 411-418. Singapore: ACM, 2008.