

Analizando o Uso de uma Grande Rede de Streaming de Vídeos na Divulgação de Conteúdo Político

Felipe Giori Guilherme dos Santos Flavio Figueiredo
Fabrício Benevenuto Jussara M. Almeida

¹Departamento de Ciência da Computação – Universidade Federal de Minas Gerais

{felipegiori, guilhermesoares, flaviovdf, fabricio, jussara}@dcc.ufmg.br

Resumo. Atualmente, o YouTube como plataforma de vídeos tem uma presença ubíqua na disseminação de conteúdo online. Desta forma, é esperado tanto que marqueteiros quanto criadores de conteúdo façam uso do YouTube para disseminação de produtos e ideias. Neste artigo, analisamos o tráfego de vídeos políticos e vídeo propagandas no YouTube durante as eleições de 2018. Nossa análise é realizada através de personas de rede sintéticas desenvolvidas de modo a simular o comportamento humano. Durante o período de aproximadamente 40 dias, simulamos eleitores de diferentes estados (através de VPNs), gênero (através de configurações de conta), além de alinhamento político (através de acessos a canais classificados como sendo de esquerda, centro e direita). Durante o experimento, as diferentes personas estavam susceptíveis às vídeo propagandas que são pareadas com os vídeos acessados de forma automatizada. Com base nos dados coletados, apresentamos um estudo focando em três aspectos: (1) características de popularidade do conteúdo dos canais; (2) na diferença de exposição de propagandas entre diferentes personas; (3) uma análise de partidos que mais fizeram uso da plataforma.

Abstract. As a streaming platform, YouTube currently has a ubiquitous online presence on disseminating video content. In this sense, both content providers, as well as marketers, make use of YouTube to promote ideas from different ideological spectra. In this paper, we investigate the use of YouTube as a political platform during the 2018 Brazilian elections. Our analysis is performed through synthetic network personas developed to simulate human behavior. During an approximate 40-day period, we simulated voters from different states, gender (through account settings), and political alignment (through access to channels ranked as left, center, and right wing). During the experiment, our personas were susceptible to video advertisements that are paired with videos accessed in an automated way. Based on the data collected, we present a study focusing on three aspects: (1) characteristics of the popularity of the content of the channels; (2) the difference in exposure of advertisements between different people; (3) an analysis of political parties that made the most use of the platform.

1. Introdução

Nós estamos na era do consumo de vídeos na web. Todos os dias, mais de 1 bilhão de horas de conteúdo em vídeo é assistido mundialmente no YouTube¹. Além disso, o YouTube é o segundo website mais acessado no Brasil, atrás apenas do Google. Apesar do

¹<https://www.youtube.com/intl/en-GB/yt/about/press/>

YouTube ser uma plataforma famosa para encontrar conteúdo voltado para o entretenimento, ele também pode ser uma plataforma para o compartilhamento de informação de vários tipos como educação, saúde e política (nosso foco).

Em um ano eleitoral no Brasil, os candidatos políticos começam suas campanhas aproximadamente 2 meses antes do início das eleições. Entretanto, antes mesmo desse período, meios de comunicação já começam a cobertura dos possíveis candidatos que concorrerão a cargos, e os eleitores são expostos e começam a consumir esse tipo de conteúdo. Uma das formas que esse consumo pode ocorrer é através dos vídeos no YouTube. Dada a importância do YouTube no mercado de *streaming* de vídeos no Brasil, além das recentes regulamentações sobre o uso de propagandas políticas na web², podemos esperar que candidatos utilizem a plataforma para alavancar suas candidaturas.

No sistema, vídeos sobre política podem ser vistos na plataforma através de conteúdo criado por usuários, ou através de propagandas que são pareadas com esses vídeos, as chamadas vídeo propagandas. Esse pareamento é realizado através de algoritmos proprietários. Como consequência, surge um novo cenário, ainda pouco estudado, em que candidatos políticos podem utilizar o sistema de vídeo propagandas do YouTube para aumentar o alcance de sua candidatura. Neste trabalho, visamos analisar como as pessoas são expostas a conteúdo no YouTube, via vídeos requisitados ou vídeo propagandas automaticamente pareadas com os primeiros, gerando tráfego correspondente, focando em um período de importante campanha eleitoral no Brasil.

Para realizar esta pesquisa, desenvolvemos um coletor, *crawler*, baseado em personas sintéticas. Tais personas consistem de códigos automatizados que fazem uso de contas do Google³ para simular pessoas comuns (potenciais eleitores). Cada persona é identificada por diferentes atributos de conta como: *gênero, preferências e localização*. Ao criar as contas do Google, definimos o gênero. Para a localização, fizemos *deploy* do código usando *Virtual Private Networks (VPNs)* em dois estados, Paraíba e Minas Gerais. Por fim, para definir as preferências, cada persona foi configurada para seguir um sub-conjunto de 75 canais que produzem conteúdo político. Os sub-conjuntos consistem de canais de esquerda, direita e centro. Desta forma, o *crawler* tenta reproduzir o comportamento de usuários da plataforma ao assistir vídeos online.

As personas executaram por um período de 40 dias englobando tanto o primeiro quanto o segundo turno das eleições de 2018. No total, 2.568 vídeos diferentes de 75 canais foram assistidos 5.545 vezes. Além do mais, ao acessar diferentes conteúdos dos canais, ou vídeo conteúdos, as personas eram susceptíveis às vídeos propagandas que são exibidas em conjunto com o conteúdo [Arantes et al. 2017]. Nossa análise foca nos seguintes aspectos: (1) propriedades do conteúdo dos canais assistidos; (2) propriedades das vídeo propagandas exibidas em conjunto com os vídeos acessados pelas personas; (3) um análise de partidos políticos que mais fizeram uso da plataforma.

Alguns dos nossos principais resultados mostram que: (1) no YouTube, vídeos de conteúdo de direita receberam mais interações positivas (e.g., likes) do que os vídeos de esquerda; (2) o YouTube foi explorado por marqueteiros para propagar conteúdo político na forma de propagandas; e (3) a maioria das propagandas foram disseminadas no pri-

²<http://www.justicaeleitoral.jus.br/arquivos/propaganda-eleitoral-na-internet>

³<https://account.google.com>

meiro turno, pois a maioria dos políticos foram candidatos derrotados ou eleitos que abandonaram a plataforma no segundo turno.

O resto deste artigo está organizado da seguinte forma. A Seção 2 apresenta os trabalhos relacionados, e a Seção 3 discute nossa metodologia de coleta de dados. A Seção 4 analisa o conteúdo dos canais em si. Em seguida, a Seção 5 foca nas vídeos propagandas, enquanto a Seção 6 foca em entender quais políticos fizeram mais uso da plataforma. Por fim, a Seção 7 conclui o trabalho.

2. Referencial Teórico e Trabalhos Relacionados

Nesta seção apresentamos uma discussão sobre trabalhos anteriores que estudaram tanto aspectos políticos em redes sociais, como também análises de dados e vídeo propagandas no geral. Antes de iniciar, definimos alguns termos que serão utilizados em todo o artigo. Ao acessar o YouTube, um **usuário**, geralmente navega pela plataforma (realizando buscas ou acessando *links*) até um **vídeo conteúdo**. No escopo deste trabalho, tal vídeo conteúdo é definido como o conteúdo de mídia final que o usuário deseja assistir, ou fazer *stream*. Vídeos conteúdos no YouTube são disponibilizados por diferentes **canais**, sendo muitos canais comumente focados em tópicos específicos: política, notícias, fofoca, etc.

Ao acessar um vídeo conteúdo no YouTube, algoritmos de leilões [Sizov and Piller 2016, Wang et al. 2014, Yeo et al. 2017, Chen et al. 2015, Mei et al. 2017, Krishnan and Sitaraman 2013] são executados para selecionar **vídeo propagandas** a serem **pareadas** com o vídeo conteúdo. Este pareamento implica em exibir a vídeo propaganda selecionada para o usuário, geralmente antes do vídeo conteúdo (embora possa também ser exibida durante ou depois do *stream* do vídeo). Estas vídeo propagandas, utilizadas para vender produtos e serviços ou divulgar ideias no sistema, têm duração variada, embora frequentemente usuários possam interromper suas exibições após alguns segundos (5 segundos) [Arantes et al. 2016]. O YouTube permite que qualquer pessoa crie uma vídeo propaganda. Dado um orçamento e o tipo de público que se quer atingir, a vídeo propagada é exibida para usuários por meio dos pareamentos com os vídeo conteúdos requisitados. Esse sistema de vídeo propagandas possui diretrizes que mudam bastante de entre diferentes países⁴.

Pesquisas similares feitas sobre eleições de 2006 e 2008 nos Estados Unidos [Ridout et al. 2010, Gueorguieva 2008] estudaram a popularidade de vídeos com conteúdo político no YouTube, analisaram como foi o uso desses vídeos nas eleições e os possíveis efeitos que esses vídeos tiveram nos resultados das eleições. Porém, seus trabalhos focaram em vídeos de campanha e vídeos virais que foram compartilhados por usuários. O diferencial de nosso trabalho é que analisamos o uso de uma ferramenta do YouTube que pode ser usada para a propagação de conteúdo político. Isto é, um usuário que busca ou acessa um vídeo que foi compartilhado com ele, tomou a decisão de ver esse conteúdo. Entretanto, antes de ter acesso ao conteúdo requisitado, ele pode ser exposto a uma vídeo propaganda e precisará assisti-la por um tempo mínimo⁵ antes de ter acesso ao vídeo requisitado⁶.

⁴<https://support.google.com/adspolicy/>

⁵Tipicamente usuários podem interromper a transmissão de uma vídeo propaganda após 5 segundos, embora em alguns casos a interrupção não seja permitida.

⁶É importante ressaltar que embora o YouTube possua um modelo de negócios baseado em propagandas,

Existe uma grande gama de trabalhos prévios que abordam o tema de propagandas online [Sizov and Piller 2016, Wang et al. 2014, Yeo et al. 2017, Chen et al. 2015, Mei et al. 2017]. Em particular no contexto de vídeo, o trabalho de Arantes *et al.* [Arantes et al. 2016, Arantes et al. 2017, Arantes et al. 2018] explorou características de tráfego da web [Arantes et al. 2016], monetização [Arantes et al. 2017], além de estudos com usuários [Arantes et al. 2018]. Todos estes estudos focaram em propagandas na forma de vídeos no YouTube. Além disso, alguns autores analisaram as propriedades de cache e tráfego de anúncios de vídeo em dispositivos móveis [Amarie et al. 2014]. O uso de personas foi inspirado no trabalho de [Carrascosa et al. 2015], que fez uso de personas sintéticas para analisar propagandas em máquinas de busca. Aqui, nós usamos de estratégia similar, mas aplicada ao contexto de *stream* de vídeos no YouTube.

3. Metodologia

Nesta seção iremos explicar como fizemos a coleta dos dados utilizando personas sintéticas de eleitores. Iniciamos na Seção 3.1 com uma discussão do nosso coletor de dados, *crawler*, baseado em personas. Em seguida, na Seção 3.2 descrevemos a base de dados coletada. Na Seção 4 iniciamos a discussão dos resultados do estudo.

3.1. Coleta de Dados Baseada em Personas

Para conseguirmos atingir o objetivo do trabalho, nós criamos coletores, ou *crawlers* através da automatização do navegador Firefox⁷ utilizando a biblioteca Selenium⁸. O crawler simulou personas baseadas em eleitores brasileiros. As personas criadas são eleitores dos gêneros masculino ou feminino, que moram nos estados de Minas Gerais ou Paraíba, têm 30 anos de idade e são dos alinhamentos políticos de esquerda, direita ou centro. Sendo nossa pesquisa realizada no estado de Minas Gerais, roteamos tráfego para o estado da Paraíba usando uma *Virtual Private Network*, ou VPN, através do serviço fornecido pela empresa VPNArea⁹. Executamos metade das personas em Minas Gerais, sem VPN, e a outra metade na Paraíba utilizando a VPN. Para cada persona criamos uma conta do Google, e as variáveis relacionadas a gênero e idade foram adicionadas às mesmas.

Para escolher as variáveis do estudo, gênero, localidade e alinhamento, nos motivamos nas explicações fornecidas pelo YouTube indicando os motivos pelo qual uma vídeo propaganda foi pareada com um vídeo conteúdo. Isso é feito através do botão “Por que estou vendo isto?” que está presente em algumas vídeo propagandas. Ao clicar no botão, uma janela é aberta e são dadas várias razões do porquê um usuário está vendo aquela propaganda. Algumas das razões listadas são gênero, hora do dia e localidade. Uma captura de tela pode ser vista na Figura 1.

Por tais motivos, escolhemos executar as personas em dois estados, Minas Gerais e Paraíba, e em três intervalos de horário, de 9h às 10h, 15h às 16h e de 19h às 20h, além de explorar os dois gêneros masculino e feminino. Além das variáveis relacionadas à conta de usuário do Google, também variamos o conteúdo que cada persona assistiu. As dividimos em três grupos de acordo com seu alinhamento político: esquerda, centro

recentemente ele apresentou o YouTube Premium, onde propagandas não são apresentadas ao usuário caso ele pague uma assinatura mensal.

⁷<http://firefox.org>

⁸<http://www.seleniumhq.org>

⁹<http://vpnarea.com>

Este anúncio baseia-se em:

- o horário do dia, o website que você estava visualizando ou sua localização geral (por exemplo, país ou cidade);
- Este anúncio foi exibido com base nas pesquisas recentes da sua Conta do Google.
- Informações na sua Conta do Google.
- O gênero que você adicionou à sua Conta do Google

Visite as [Configurações de anúncios](#) do Google para saber mais sobre como os anúncios são segmentados ou para desativar os anúncios personalizados.

FECHAR

Figura 1. Janela ao clicar no botão: "Por que estou vendo isto?"

Tabela 1. Canais com conteúdo político

Esquerda			Centro			Direita		
Canal	# Vis.	Víd.	Canal	# Vis	Víd.	Canal	# Vis.	Víd.
Lula	1.1M	145	Social Good Brasil	200k	588	Central de Oposição	69M	94
Marina Silva	5M	929	Lidice da Mata	166k	865	Notícias Políticas BR	3M	36
Jean Wyllys	5.6M	599	Conectas Dir. Humanos	89k	186	Primeira Mão	91M	1915
PT	6.1M	1613	Católicas direito d.d.	377k	91	Bolsonaro TV	80M	508
Jandira Feghali	352k	428	Fund. Amazonas Sust.	48k	213	Mamaefalei	218M	588
Luciana Genro	1.7M	237	Intervozes	69k	85	José Marcio Art.142	34M	1070
Marcelo Freixo	4.6M	1455	MAB Comunicação	234k	213	conversaTUBE	58M	407
Gleisi Hoffmann	120k	145	GIFE	169k	269	Canal do Opressor	138M	379
PSOL 50	120k	112	Articulação Semiárido	2.4M	184	João Doria	1.6M	937
Chico Alencar 500	58K	68	Instituto Ethos	113k	329	Sargento Moterani	1.8M	222
Anistia Int. Brasil	1.1M	237	Fund. Perseu Abramo	928k	1220	Jair Bolsonaro	80M	1174
Manuela d'Ávila	1.8M	634	Joice Hasselmann TV	183M	1662	Sargento Fahur	19M	211
PT na Câmara	660k	2020	Roda Viva	51M	1620	MBL	147M	1524
PSDTV	1M	3937	Rodrigo Constantino	5.5M	463	PSDB	2.3M	1651
Kátia Abreu	400k	310	Leandro Ruschel	4.3M	283	NOVO 30	1.3M	200
CUT Brasil	154k	404	Ideias Radicais	81M	1898	Eduardo Bolsonaro	14M	837
Pablo Villaça	4.5M	381	Leandro Karnal - Adm.	14M	983	Alvaro Dias	1.7M	3001
Vanessa Grazziotin	256k	1129	Canal Brasil	62M	1240	Coronel Telhada	2.3M	1207
Dilma Resistente	4M	588	Jornalismo TV Cultura	32M	16524	Ana Amélia Lemos	388k	1986
PSTU	2.3M	943	Rádio BandNews FM	43M	2124	ACM Neto	890k	366
TV 247	81M	5352	Uma Gota no Oceano	1.9M	198	Inst. Mises Brasil	865k	676
C. A. c/ Paulo H. A.	92M	1402	institutoakatu	1M	184	Flavio Bolsonaro	7.3M	507
Carta Capital	10M	1076	Agência Pública	914k	279	Canal do Otário	81M	815
D. Centro do Mundo	20M	1081	Inst. Socioambiental	1.1M	297	CanalBolsonaro2018	31M	2899
Saia da Matrix	1.5M	221	Repórter Brasil	1M	196	Pesadelo de Q. Pol.	1.1M	207

ou direita. As personas então se inscreveram e acompanharam diariamente 25 canais do YouTube, de acordo com seu alinhamento político. Foram selecionados apenas canais que tinham uma frequência de *upload* de pelo menos 1 vídeo por mês, pois canais que não são muito ativos tinham grande chance de não serem vistos.

Os canais selecionados podem ser vistos na Tabela 1. A categorização de canais por viés político foi definida inicialmente com base no sítio da web Monitor do Debate Político no Meio Digital¹⁰, um projeto do Grupo de Pesquisa em Políticas Públicas para o Acesso a Informação da USP. A listagem do sítio contém os nomes de contas famosas da Internet categorizadas por viés. Para termos o mesmo número de canais por viés, também realizados buscas nos Google e YouTube por canais populares de cada viés, chegando assim na lista final de canais apresentada na Tabela 1.

Com as personas criadas e inscritas nos canais de seu respectivo alinhamento político, as mesmas acessam o YouTube diariamente e assistem aos vídeos recentemente

¹⁰<http://www.monitordigital.org/>

lançados pelos canais inscritos. Limitamos essa janela de tempo para um dia, ou seja, as personas assistem apenas aos vídeos lançados no dia que elas estão navegando. Caso uma persona não tenha vídeos para assistir, porque nenhum dos 25 canais lançou algum vídeo, a persona não assiste nada e encerra sua sessão. Também, é possível que uma persona veja um mesmo vídeo entre sessões, mas nunca o mesmo vídeo na mesma sessão.

Os crawlers foram executados em ambiente Linux e utilizando o Docker para isolar as mesmas¹¹. Foram criados quatro *containers* e cada *container* executava três personas, uma para cada alinhamento político, do mesmo estado e mesmo gênero. Dentro dos dockers referentes às personas da Paraíba, executávamos um script para ativar a VPN. Desta forma, cada container tinha seu próprio IP e pudemos simular a localização através da VPN. Nós dividimos a execução das personas em dois períodos referentes ao primeiro e segundo turnos das eleições de 2018. A primeira execução ocorreu do dia 19/09/2018 ao dia 07/10/2018, data do primeiro turno. A segunda execução foi feita de 08/10/2018 a 28/10/2018. No primeiro turno apenas as personas da Paraíba foram executadas. No segundo turno as personas dos dois estados foram executadas.

As vídeo propagandas associadas aos vídeos assistidos foram coletadas através de requisições HTTPS. Conseguimos detectar um tipo de requisição que possui o id do vídeo e o id da vídeo propaganda, e isolamos essas requisições para fazer o pareamento. Após o término do experimento, também coletamos meta-dados dos diferentes vídeo conteúdos e vídeo propagandas coletados. Esta coleta fez uso da API do YouTube¹². Em particular, coletamos o número de visualizações global, isto é, o número de streams fornecido pelo YouTube. Além das visualizações, também coletamos o número de curtidas e não-curtidas (*likes e dislikes*), além do total de comentários recebidos por um vídeo. Tais dados complementam a nossa visão das personas que são focadas em um único usuário. As estatísticas globais mostram o sucesso dos vídeos como um todo. A coleta dos dados globais foi feita na semana do dia 17 até o dia 21 de Dezembro. Assim, as mesmas capturam estatísticas que vão além do período das eleições. De qualquer forma, nossas análises focam nos vídeos que foram propagados durante tal período.

3.2. Descrição dos dados

A Tabela 2 sintetiza os número das exibições para cada persona e para cada turno. Através da tabela, observe que cada persona fez stream de pelo 122 vídeos únicos. Similarmente, pelo menos 53 propagandas foram exibidas para cada persona. Ao todo foram realizadas 5.545 exibições, *streams*, de vídeos, sendo 2.606 no primeiro turno e 2.939 no segundo turno. Com relação ao número de vídeo propagandas, foram 2.050 exibições no total, sendo 910 exibições no primeiro turno e 1.140 exibições no segundo turno.

Note pela tabela que o número de exibições tanto de vídeo conteúdo quanto de vídeo propagandas é similar entre as personas. Tal resultado é esperado dado que cada persona acessa aproximadamente o mesmo número de canais em intervalos de tamanho similares. A maior diferença existente é entre o primeiro e segundo turno para as personas da Paraíba. Nestas, podemos ver uma queda de propagandas no segundo turno possivelmente devido ao fato de que muitos candidatos já foram eleitos: deputados, senadores, além de governadores já eleitos no primeiro turno.

¹¹<http://docker.org>

¹²<https://developers.google.com/youtube/v3/>

Tabela 2. Visão geral das personas e das visualizações

Primeiro turno					
Estado	Gênero	Alinhamento Político	# exibições (conteúdo)	# exibições (propagandas)	
PB	Masculino	Direita	505	140	
PB	Feminino	Direita	565	96	
PB	Masculino	Esquerda	482	236	
PB	Feminino	Esquerda	487	160	
PB	Masculino	Centro	268	134	
PB	Feminino	Centro	299	144	
Segundo turno					
PB	Masculino	Direita	332	76	
PB	Feminino	Direita	356	53	
PB	Masculino	Esquerda	254	148	
PB	Feminino	Esquerda	268	116	
PB	Masculino	Centro	206	95	
PB	Feminino	Centro	174	58	
MG	Masculino	Direita	327	139	
MG	Feminino	Direita	219	66	
MG	Masculino	Esquerda	293	138	
MG	Feminino	Esquerda	185	87	
MG	Masculino	Centro	203	109	
MG	Feminino	Centro	122	55	

Tabela 3. Top 5 canais mais assistidos por alinhamento político

Esquerda		Centro		Direita	
Canal	# Vis.	Canal	# Vis.	Canal	# Vis.
TV 247	961	Rádio BandNews FM	333	Primeira Mão	386
Diário do Centro do Mundo	270	Jornalismo TV Cultura	233	Canal Bolsonaro 2018	276
Dilma Resistente	96	Joice Hasselmann TV	154	Jair Bolsonaro	269
PT - Partido dos Trabalhadores	92	Canal Brasil	122	João Dória	261
Manuela d'Ávila	66	Ideias Radicais	67	Álvaro Dias	166

4. Análise dos Canais

Durante a execução as personas com alinhamento político de direita assistiram a 2.304 vídeos, o que corresponde a 41,55% do total, as de esquerda assistiram a 1.969 vídeos, o que corresponde a 35,51% do total e as de centro assistiram a 1.272 vídeos, o que corresponde a 22,94% do total. Os 5 canais mais assistidos de cada alinhamento político podem ser vistos na Tabela 3 com o número de visualizações para cada canal.

Em nosso experimento, as personas com alinhamento político de direita viram mais vídeos, seguidos pela de esquerda e depois centro. Além disso, no dia 14/12 coletamos dados relacionados a popularidade dos vídeos assistidos na época do experimento, a fim de visualizar o engajamento do público nesses vídeos. Na data da coleta dos dados 166 vídeos haviam sido deletados e não foi possível conseguir suas informações.

A Figura 2 mostra o complemento da função distribuição acumulada (CCDF) para diferentes métricas de um vídeo conteúdo (visualizações, likes/dislikes e comentários). Na figura, as métricas like, dislike e comentários foram normalizadas pelo número de visualizações, pois é esperado que esses números estejam diretamente relacionados ao número de visualizações. Iniciamos nossa exploração com a CCDF da Figura 2a, referente ao número de visualizações. Os vídeos assistidos, independentemente do alinhamento político, obtiveram uma quantidade de visualizações bem similar até a faixa de 10^4 visualizações. Após essa faixa, os vídeos de direita conseguem manter um número médio de visualizações maior. No geral, a mediana de visualizações que os vídeos vistos

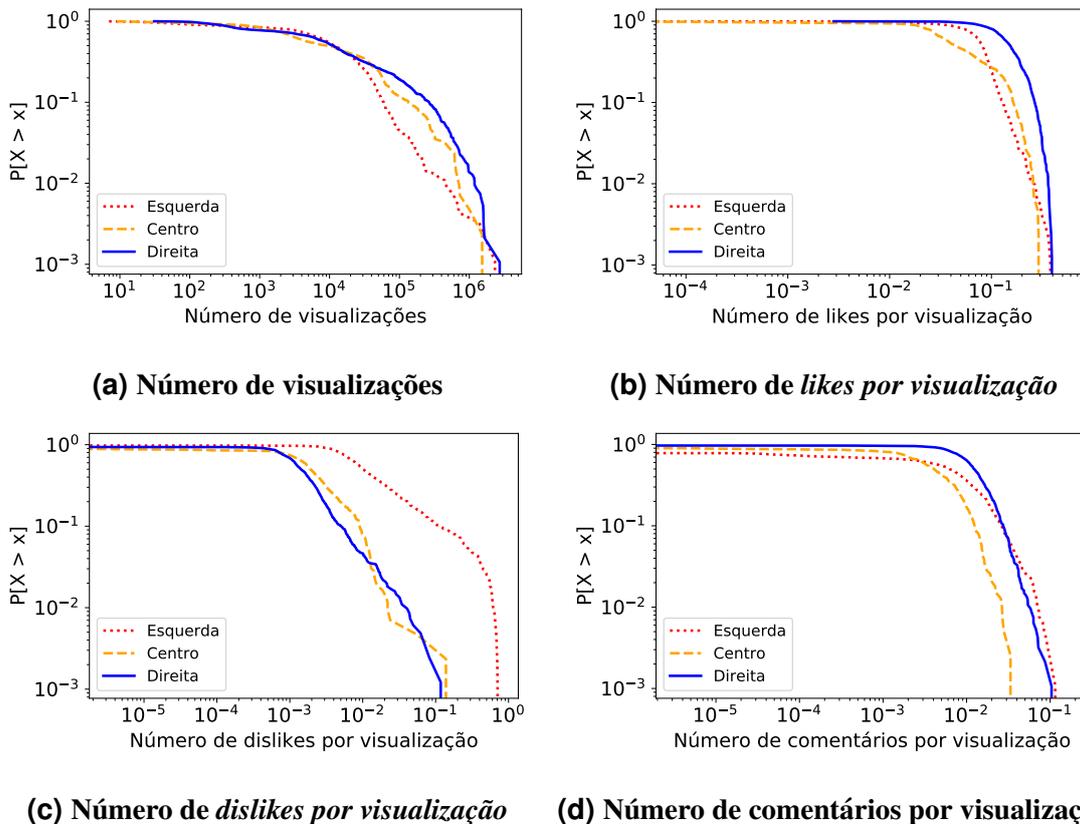


Figura 2. Distribuições de Probabilidades para Métricas de Vídeo Conteúdos

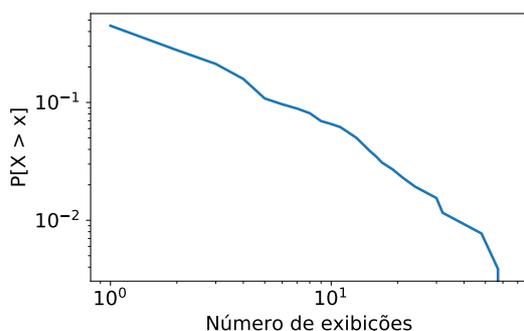
no experimento atingiram foi de 10.474 visualizações para os vídeos de esquerda, 1.889 para os vídeos de centro e 11.387 para os vídeos de direita. Analisando pelo 75º percentil, esses números vão para 29.884 visualizações para a esquerda, 48.882 para o centro e 57.859 para a direita.

Os vídeos de direita não só foram os mais vistos no geral, como também foram os que atraíram, em termos relativos, uma quantidade maior de engajamento positivo dos usuários. Por exemplo, o 25% dos dados (terceiro quartil) do número normalizado de likes por visualização foi de pelo menos 0.1 para os vídeos de esquerda, 0.11 para os de centro e 0.2 para os de direita. Já com relação ao número de *dislikes*, o mesmo quartil do número normalizado de dislikes por visualização foi de 0.03 para esquerda, 0.004 para o centro e 0.002 para a direita. As CCDFs das Figuras 2b e 2c são referentes ao número de *likes* e *dislikes*. Pelas CCDFs é possível ver que os vídeos de centro e direita possuem bem menor rejeição que os vídeos de esquerda. Considerando comentários, Figura 2d, os vídeos de direita, em mediana, possuem mais comentários. Contudo não analisamos se os comentários são positivos ou negativos. Tal análise é deixada para trabalhos futuros.

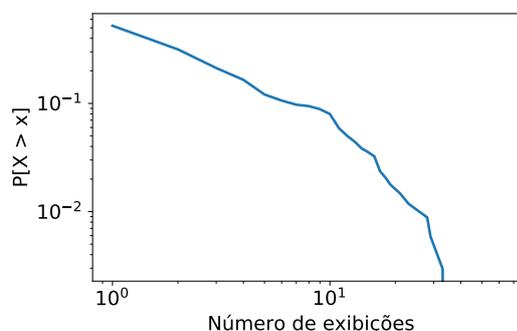
Os resultados até agora mostram como foram as interações dos usuários com conteúdo político no YouTube. Em particular, analisamos variáveis de interações como: visualizações, likes, dislikes e comentários. É interessante que canais de direita recebem mais interações positivas do que aqueles de esquerda. Além dos mais, tais canais recebem marginalmente mais visualizações. Tal resultado indica que possivelmente este alinhamento político obteve “mais sucesso” na rede, ou que o cenário político do país re-

Tabela 4. Tabelas de pareamento separadas por alinhamento político

	Esquerda			Centro			Direita		
	Homem	Mulher	Total	Homem	Mulher	Total	Homem	Mulher	Total
MG	47,1%	47,03%	47,07%	53,7%	45,08%	50,46%	42,51%	30,14%	37,55%
PB	52,17%	36,56%	44,27%	48,31%	42,71%	45,51%	25,81%	16,18%	20,76%
Total	50,73%	38,62%	44,94%	49,92%	43,19%	46,78%	30,5%	18,86%	24,74%



(a) CCDF - Primeiro turno



(b) CCDF - Segundo turno

Figura 3. CCDFs das exibições de vídeo propagandas para os dois turnos

fletido no YouTube tende para tal viés. Na próxima seção vamos analisar as propagandas e partidos que fizeram uso da rede.

5. Pareamentos e Vídeo Propagandas

Nesta seção, realizamos a nossa análise das vídeo propagandas pareadas com os vídeos conteúdos. No geral, 36,97% dos vídeos tiveram uma vídeo propaganda pareada. Separando pelos turnos, esse valor fica 34,92% para o primeiro turno e 38,79% para o segundo turno. Separando pelos alinhamentos políticos das personas, esse valor é 44,94% para a esquerda, 46,78% para o centro e 24,74% para a direita. Ou seja, no nosso experimento, as personas da direita, em média, viram menos propagandas.

Para aprofundar a análise, fizemos tabelas separadas por cada alinhamento político, utilizando os dados dos dois turnos, então, calculamos o valor do número médio de vídeo propagandas exibidas por vídeo para os gêneros e os estados. A Tabela 4 mostra a porcentagem de pareamentos correspondente a essas variáveis para cada alinhamento político. Neste caso, por exemplo, uma persona que é homem, possui alinhamento político de direita e mora em Minas Gerais viu propagandas em 42,51% dos seus vídeos.

Além das frações de pareamentos, também analisamos a popularidade das vídeo propagandas pareadas para as nossas personas. Portanto, os gráficos da Figura 3 apresentam as CCDFs do número de exibições das propagandas para os dois turnos. Olhando as figuras, inicialmente percebemos que a maioria das propagandas foram exibidas no máximo uma ou duas vezes. No primeiro turno, a mediana das exibições das vídeo propagandas foi igual a 1, ou seja, pelo menos metade das vídeo propagandas pareadas foi exibida apenas uma vez. O número máximo de exibições de uma vídeo propaganda, no primeiro turno, foi de 63 exibições. Esse mesmo tipo de comportamento também pode ser visto no segundo turno. A mediana das exibições das vídeo propagandas foi 2 e o número máximo de exibições de uma vídeo propaganda foi de 67 exibições. Para concluir, observando as CCDFs é possível perceber uma tendência onde poucas propagandas

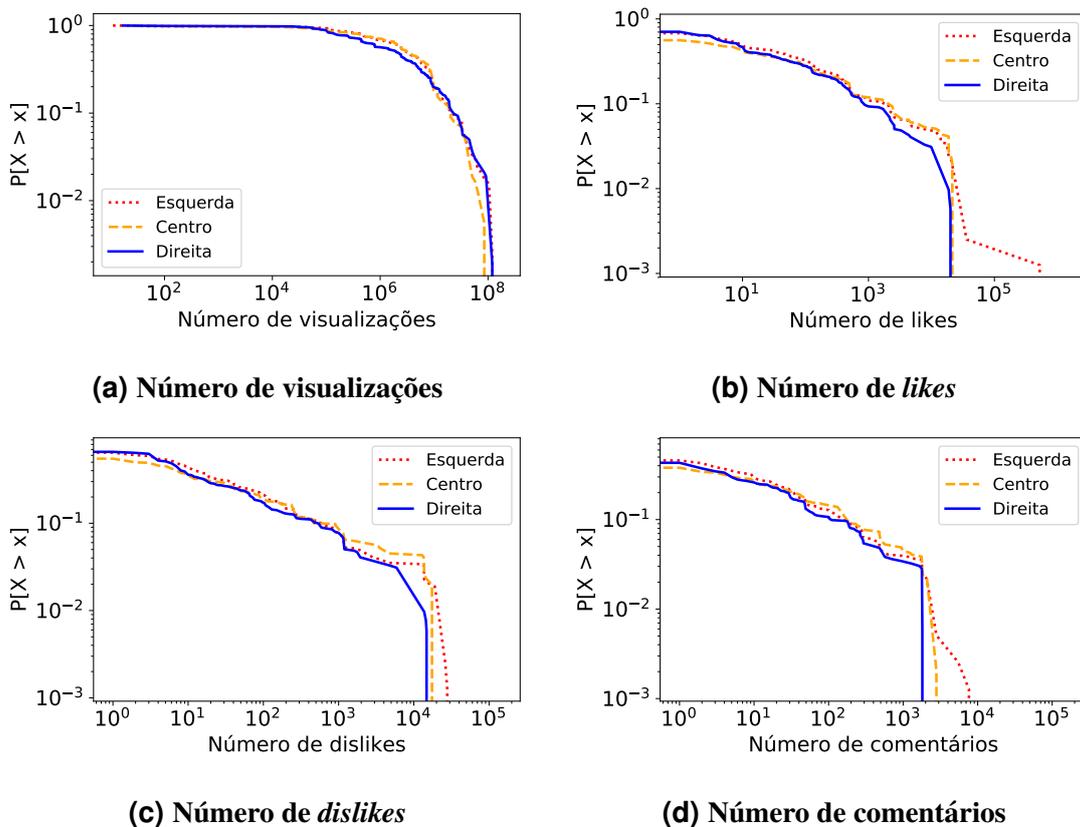


Figura 4. Distribuições de Probabilidades para Métricas de Vídeo Propagandas

são exibidas muitas vezes e muitas propagandas são exibidas poucas vezes. Por exemplo, para ambos os turnos, mais de 90% das propagandas possuem menos de 10 exibições.

Por fim, a Figura 4 apresentamos as distribuições das métricas de popularidade (visualizações, comentários, likes e dislikes) para as vídeo propagandas. Diferente do ocorrido com os vídeo conteúdos, analisando os gráficos percebe-se que o número de cada métrica é bastante similar através dos diferentes perfis políticos. Embora tal comportamento seja similar, na próxima seção vamos aprofundar nossa análise nas propagandas realizando uma análise do conteúdo das mesmas. Vamos mostrar como alguns candidatos fizeram uso da plataforma para divulgação de conteúdo político, além de descrever as principais marcas que potencialmente realizaram algum tipo de propaganda.

6. Análise de Partidos e Políticas

Até o momento, o nosso estudo focou apenas em questões de popularidade no âmbito local, através das personas, além do global, através dos dados da API do YouTube, dos vídeo conteúdos e vídeo propagandas. Nesta seção, analisaremos as propagandas em mais detalhes focando nos produtos, ideias e partidos que fizeram uso do YouTube para disseminação de informação durante o período das eleições.

A fim de caracterizar o conteúdo das vídeo propagandas, elaboramos um questionário que tem como objetivo extrair informações que permitam agrupar as propagandas. As perguntas do questionário podem ser vistas na Tabela 5. Inicialmente, um pesquisado, autor do artigo, respondeu ao questionário para as 50 vídeo propagandas mais po-

Tabela 5. Perguntas do questionário de caracterização de vídeo propagandas

Pergunta	Tipo
Qual a marca que está anunciando?	Aberta
Qual o produto sendo anunciado	Aberta
Qual a categoria da vídeo propaganda?	Fechada
Esta vídeo propaganda é referente à campanha de um candidato?	Sim/Não

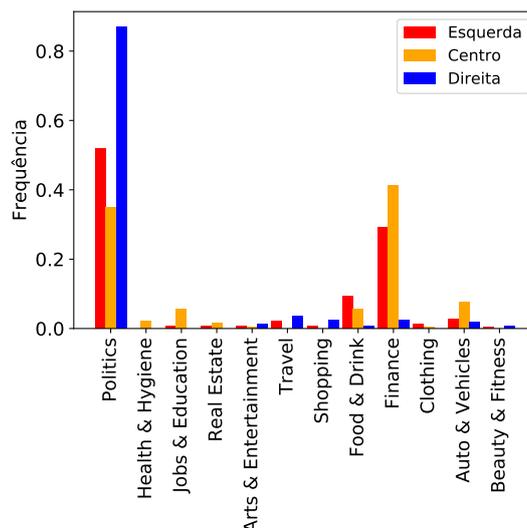


Figura 5. Distribuição das categorias por alinhamento político - Primeiro turno

pulares do primeiro turno. Após o término das respostas, um outro pesquisador (também co-autor) verificou as respostas do primeiro atingindo um nível de 96% de concordância. Após averiguar tal concordância, foi feita a codificação de 80 vídeo propagandas do segundo turno. Os 130 vídeos codificados foram os mais populares e cobrem a maioria, 65%, das exibições das personas. Além da codificação, também exploramos meta-dados das propagandas como os canais que fizeram o envio da mesma no sistema para entender o conteúdo divulgado durante as eleições.

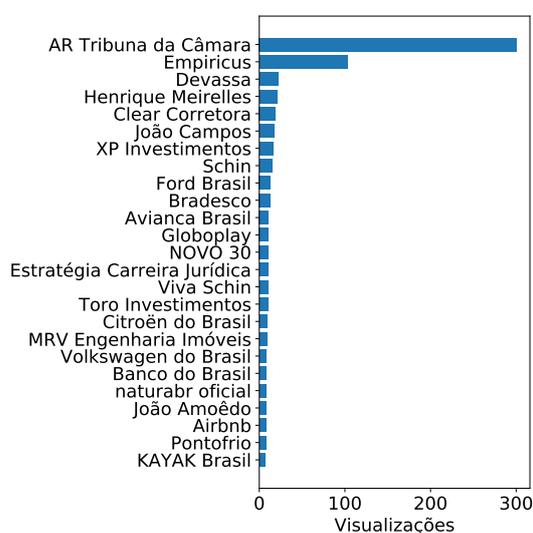
No primeiro turno, 259 propagandas foram exibidas 910 vezes. As 50 propagandas mais exibidas foram visualizadas 603 vezes, ou 66,26% do total de exibições. Dessas 603 exibições, 247 foram referentes a personas da esquerda, 194 do centro e 162 da direita. A distribuição das categorias para cada alinhamento pode ser visto na Figura 5.

Através do gráfico é possível perceber que as personas tenderam a ser mais expostas a vídeo propagandas de cunho político, principalmente as personas que assistiram vídeos da direita política. Das 162 visualizações de propagandas que as personas da direita tiveram, 141 foram de cunho político, ou 87%. A categoria de finanças, que engloba principalmente propagandas relacionadas a serviços bancários e o mercado financeiro, foi bastante exibida para as personas da esquerda e centro. Para uma análise mais aprofundada, selecionamos as dez marcas mais exibidas isolando cada variável de interesse.

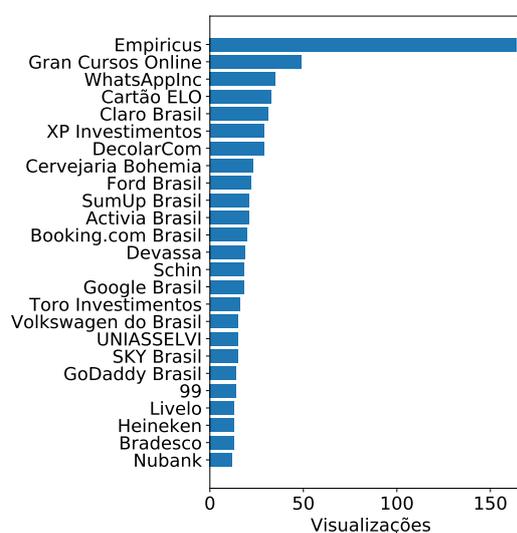
A Tabela 6 apresenta as dez marcas mais exibidas para as personas por alinhamento político e gênero. Na tabela, apresentamos não só o ranking das principais marcas, como também o percentual das vídeo propagandas codificadas onde aquela marca aparece. Note que no primeiro turno, diferentes políticos fizeram uso do YouTube para

Tabela 6. Marcas mais exibidas por alinhamento político - 1º turno

Esquerda		Centro		Direita	
Marca	Frequência (%)	Marca	Frequência (%)	Marca	Frequência (%)
André Regis	43,32	André Regis	30,93	André Regis	78,4
Empiricus	14,98	Empiricus	28,87	Henrique Meirelles	3,7
Devassa	6,48	Ford	5,16	Avianca	3,7
Clear Corretora	4,45	Estratégia Carreira Jurídica	4,12	Pontofrio	2,47
XP Investimentos	3,64	Clear Corretora	4,12	João Campos	2,47
Henrique Meirelles	3,24	XP Investimentos	3,61	João Amoêdo	1,85
João Campos	2,83	Bradesco	3,1	Citroen	1,85
Bradesco	2,83	Devassa	2,58	Globo	1,24
Avianca	2,02	Enterogermina	2,06	Natura	1,24
João Amoêdo	2,02	Henrique Meirelles	1,55	Nubank	1,24



(a) Canais primeiro turno



(b) Canais segundo turno

Figura 6. Canais que exibiram mais propagandas

marketing. Em particular, temos o André Regis e João Campos, ambos foram candidatos a deputado federal no estado de Pernambuco. O aparecimento de tais candidatos é interessante dado que o servidor da VPN estava executando em outro estado, a Paraíba. Tal resultado mostra evidências de que em um cenário político o direcionamento de propagandas por local feito pelo sistema talvez não consiga diferenciar alguns estados do Brasil. Isto é, possivelmente o mesmo deve focar em macro-regiões. Além do André Regis, vários candidatos como o João Amoedo e Henrique Meirelles. É interessante perceber como os mesmos focaram tanto em eleitores de Esquerda como de Direita, porém, ambos são menos presentes nas personas de alinhamento de centro.

Complementando nossa codificação, na Figura 6a apresentamos os canais das marcas que mais exibiram propagandas no primeiro turno, incluindo propagandas não codificadas. É possível perceber que a ordem das marcas é bem parecida com as codificadas na tabela 6. Além disso, também é possível perceber como as propagandas do André Regis (AR Tribuna na Câmara) e da Empiricus foram muito mais frequentes que as outras. No primeiro turno, esses dois canais foram responsáveis por 44% do total de visualizações de propagandas, sendo 31% correspondente ao André Regis e 11% à Empiricus.

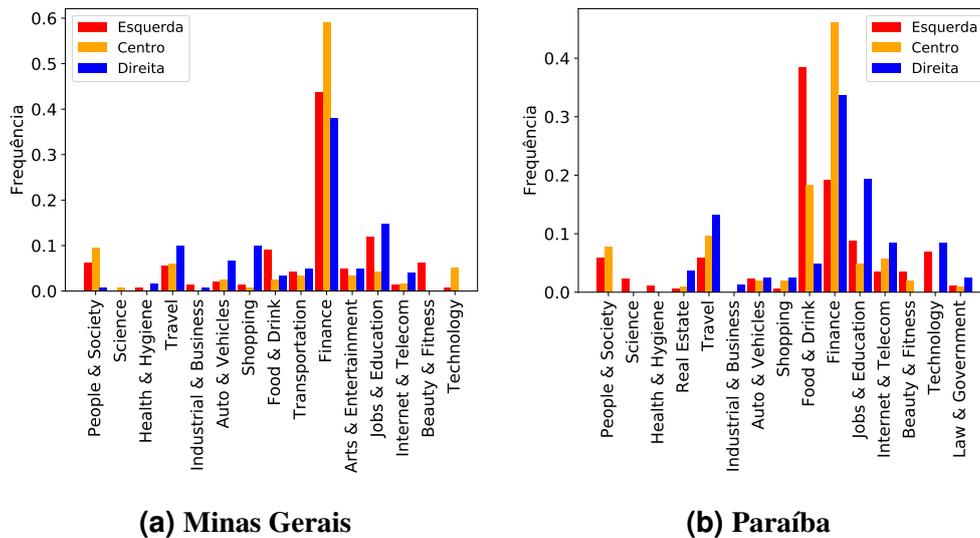


Figura 7. Distribuição das categorias por alinhamento político

Já no segundo turno, foram 339 vídeo propagandas exibidas 1140 vezes. As 80 propagandas mais exibidas foram visualizadas 737 vezes, ou 64,65% do total de exibições. Dessas 737 exibições, 314 foram referentes a personas da esquerda, 219 do centro e 204 da direita, para ambos os estados. A distribuição das categorias para cada alinhamento político e para cada estado pode ser visto nos gráficos das Figuras 7a e 7b.

A primeira categoria que nos chama a atenção é a política. Essa categoria que foi a maior no primeiro turno, agora não teve nenhuma ocorrência. Esperávamos uma queda na quantidade de propagandas, pois a maioria dos cargos políticos já havia sido preenchido no primeiro turno, porém, não esperávamos que fosse nula. A categoria relacionada a finanças foi bastante exibida em ambos os estados e para todos os alinhamentos. Na Figura 6b podemos ver as marcas e canais que mais veicularam propagandas para as personas. Embora haver propagandas que tocaram em temas relacionados ao cenário político, estas não foram classificadas como políticas. Um exemplo é o WhatsApp, que foi uma propaganda falando sobre a propagação de *Fake News* e como identificar as mesmas. Tal propaganda surgiu após o WhatsApp¹³ se tornar central no cenário político.

7. Conclusões

Neste artigo, analisamos a divulgação de conteúdo político no contexto das eleições brasileiras através da maior rede de streaming de vídeos atualmente, o YouTube. Fazendo uso de uma coleta de dados baseada em personas sintéticas de rede, foi possível coletar dados tanto de vídeo conteúdos quanto de vídeo propagandas disponibilizadas no YouTube.

Os nossos resultados mostram que o YouTube foi explorado como plataforma de propagação de informação política através da disseminação tanto de vídeo conteúdos quanto de vídeo propagandas. Este segundo tipo, as propagandas, são particularmente interessante pois as mesmas são exibidas para usuários sem a sua escolha. É importante mencionar que embora o nosso trabalho mostre que a exploração de propagandas foi pequena, os vídeos que não foram exibidos na forma de propaganda (vídeo conteúdos)

¹³<https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contra-o-pt-pelo-whatsapp.shtml>

chegam a ter milhões de visualizações. Em contrapartida, a exploração de propagandas na nossa coleta de dados foi mais ativa no primeiro turno. No segundo turno, observamos uma queda na exploração da rede, possivelmente indicando um maior foco dos marqueteiros em outras redes, como o WhatsApp. Portanto, trabalhos futuros incluem o estudo de outras redes, além de uma análise detalhada do conteúdo difundido.

Agradecimentos

Este artigo foi parcialmente financiado pelo projeto ATMOSPHERE (atmosphere-eubrazil.eu), uma cooperação entre o Ministério de Ciência Tecnologia e Inovação (número do Projeto 51119 - MCTI/RNP) com a Comissão Europeia no programa Horizon 2020 (número do *grant* 777154). Além do mais, os autores foram financiados por bolsas individuais do CNPq, Capes e Fapemig.

Referências

- Amarie, M. C., Pefkianakis, I., and Lundgren, H. (2014). Towards understanding mobile video ads. In *Ubicomp*.
- Arantes, M., Figueiredo, F., and Almeida, J. M. (2016). Understanding video-ad consumption on youtube: a measurement study on user behavior, popularity, and content properties. In *Websci*.
- Arantes, M., Figueiredo, F., and Almeida, J. M. (2017). Towards understanding the consumption of video-ads on youtube. *The Journal of Web Science*, 3(1).
- Arantes, M., Figueiredo, F., Almeida, J. M., and Prates, R. O. (2018). An investigation of user actions and experiences when exposed to youtube video ads. In *Webmedia*.
- Carrascosa, J. M., Mikians, J., Cuevas, R., Erramilli, V., and Laoutaris, N. (2015). I always feel like somebody's watching me: measuring online behavioural advertising. In *CONEXT*.
- Chen, J., Haber, E. M., Kang, R., Hsieh, G., and Mahmud, J. (2015). Making use of derived personality: The case of social media ad targeting. In *ICWSM*.
- Gueorguieva, V. (2008). Voters, myspace, and youtube: The impact of alternative communication channels on the 2006 election cycle and beyond. *Social Science Computer Review*, 26(3).
- Krishnan, S. S. and Sitaraman, R. K. (2013). Understanding the effectiveness of video ads: a measurement study. In *IMC*.
- Mei, Y., Zhao, W., and Yang, J. (2017). Maximizing the effectiveness of advertising campaigns on twitter. In *IEEE Big Data*.
- Ridout, T. N., Fowler, E. F., and Branstetter, J. (2010). Political advertising in the 21st century: The rise of the youtube ad. In *APSA 2010 Annual Meeting Paper*. APSA.
- Sizov, S. and Piller, S. (2016). Economic value of web keyword advertising campaigns. In *Websci*.
- Wang, P., Liu, Y., Meytlis, M., Tsao, H.-Y., Yang, J., and Huang, P. (2014). An efficient framework for online advertising effectiveness measurement and comparison. In *WSDM*.
- Yeo, J., Kim, S., Koh, E., Hwang, S.-w., and Lipka, N. (2017). Predicting online purchase conversion for retargeting. In *WSDM*.