

# “I just hated it and I want my money back”: Data-driven Understanding of Mobile VPN Service Switching Preferences in The Wild

Rohit Raj  
IIT Kharagpur  
rrohit2901@gmail.com

Mridul Newar  
IIT Kharagpur  
mridulnewar2000@gmail.com

Mainack Mondal  
IIT Kharagpur  
mainack@cse.iitkgp.ac.in

## Abstract

Virtual Private Networks (VPNs) are a crucial Privacy-Enhancing Technology (PET) leveraged by millions of users and catered by multiple VPN providers worldwide; thus, understanding the user preferences for the choice of VPN apps should be of importance and interest to the security community. To that end, prior studies looked into the usage, awareness and adoption of VPN users and the perceptions of providers. However, no study so far has looked into the user preferences and underlying reasons for switching among VPN providers and identified features that presumably enhance users’ VPN experience. This work aims to bridge this gap and shed light on the underlying factors that drive existing users when they switch from one VPN to another. In this work, we analyzed over 1.3 million reviews from 20 leading VPN apps, identifying 1,305 explicit mentions and intents to switch. Our NLP-based analysis unveiled distinct clusters of factors motivating users to switch. An examination of 376 blogs from six popular VPN recommendation sites revealed biases in the content, and we found ignorance towards user preferences. We conclude by identifying the key implications of our work for different stakeholders. The data and code for this work is available at <https://github.com/Mainack/switch-vpn-datacode-sec24>.

## 1 Introduction

Virtual Private Networks (VPNs) are a Privacy Enhancing Technology (PET) that has become increasingly popular over the last two decades. With the advent of commercial VPN services, VPNs are adopted by millions of users [9]. Today, VPN usage ranges from accessing geo-restricted content to accessing the web securely through public wifi [1, 15].

Prior work looked into the end-user’s awareness and usage of VPNs [43, 54]. They found that end-users have varied

expectations from VPN services. People who are emotionally invested in protecting their privacy are generally resilient VPN users. Some users might expect that a VPN can help them access Netflix US while they are on vacation in China, and they may not care much about privacy, while others might want to use a VPN to securely connect to their offices, protecting them from potential privacy threats. Intuitively, this variation in expectation, combined with massive growth in the user base and ecosystem with multiple key players [17], provides users with the choice to use the VPN service that best caters to their needs. In fact, VPN users today have the obvious option of switching between VPN providers<sup>1</sup>.

However, a user who is adopting a VPN and a user who is switching between VPN service providers are potentially different in terms of enhanced awareness resulting from previous VPN usage (and perhaps due to facing associated problems). Thus, recent work has reviewed and investigated the technical aspects of VPN infrastructure (e.g., for finding vulnerabilities) and the adoption of VPNs [43, 54]. However, it is unclear whether users switch between VPN providers and what are their key motivators (e.g., specific VPN usage preferences). This issue is further complicated due to the unclear impact of VPN-review websites on this VPN-switching behaviour. According to recent work, most of the users consider the VPN review blogs available online to choose from a large number of VPNs during VPN adoption [50]. These prior works also noted that these websites could be biased towards particular VPN service providers. However, it is unclear how these websites might cater to a more experienced VPN user who wants to switch between VPNs. This question of the requirements of existing VPN users and whether the current VPN apps and VPN review blogs cater to those desired preferences has serious implications for VPN providers, VPN-review blogs, and the general research community. Unfortunately, there is not much work focusing on understanding end-users’ behaviour and preferences regarding VPN switching. Thus, in this work, we fill this gap and investigate VPN-switching behaviour and

An extended version of this paper that includes the appendices is available for interested readers at [arXiv](https://arxiv.org/abs/2008.08881). All references to Appendix in this paper refers to the Appendix included in that [arXiv](https://arxiv.org/abs/2008.08881) version.

<sup>1</sup>Since VPNs are often used as applications (or apps), in this work, we use VPN, VPN-provider and VPN-app interchangeably.

preferences of end-users in the wild.

We focus on mobile VPN apps (due to the ubiquity of smartphones and additional metadata like reviews and # downloads provided by the app stores). We systematically selected 20 popular VPN apps on the Android and iOS platforms. Then, we collected 1,333,160 user-generated reviews for these apps from the Google Play Store and Apple App Store. Next, we created a systematic keyword search, manual labelling and automated text-classification-driven approach<sup>2</sup> to identify 185,399 reviews where users intended to switch between VPN providers (often with reason). Next, we took a deep dive into the why question—our in-depth qualitative analysis of these reviews uncovered a hierarchy of reasons for VPN-switching—we synthesized these reasons into the most desirable and undesirable properties of VPNs as indicated by users in the wild. Furthermore, we demonstrate that these reasons for VPN switching have interesting differences from prior work on VPN adoption reasons. Finally, we collected 376 VPN-review blogs and checked in detail if they potentially covered user-desired features of VPNs during switching. Specifically, we answer five key research questions.

**RQ1** *Do users switch between VPNs? Is the switching uniform across different VPN providers?*

**RQ2** *What are the key reasons driving VPN-switching?*

**RQ3** *Are the features desired during VPN-switching, by existing VPN users, different from features desired during VPN adoption?*

**RQ4** *What are the concrete features desired by users during VPN-switching?*

**RQ5** *Does information by VPN review websites align with desired requirements of users during switching VPNs?*

During our endeavour to answer the above-listed questions, we found that VPN switching is a very common phenomenon among users, and there is a clear hierarchy of factors that primarily contribute to user switching between VPNs. We also found that users have different preferences while *choosing* VPNs vs. *switching* between VPNs. This can be clearly attributed to the fact that with the usage of VPN, users' expectations and requirements change, and hence, the factors they consider. Thus, we tried to enlist and group users based on factors they primarily consider while deciding to switch VPNs. Finally, to complete the study of user switching, we looked into primary sources which users refer to aid in their VPN switching decisions, i.e. VPN recommendation blogs and found that there are multiple ways in which VPN review websites fail to address the themes which are most relevant to users when switching between VPNs. In line with earlier literature, we also found that these websites show bias towards particular VPN providers and misguide the users in their search for the right VPN service (Section 10).

<sup>2</sup>The dataset and code for this work can be found at <https://github.com/Mainack/switch-vpn-datacode-sec24>.

**Limitations.** One of our study's key limitations is the potential of our approach to miss users who indeed switched VPNs but did not write a review. Furthermore, we used keyword search to identify reviews relevant to user switching context for manual analysis, but during this, we will miss some reviews due to the limited number of VPN names used as keywords and not using terms like "switching" for the search, as that might lead to a large number of false positives making the manual process more tedious. However, this limitation won't affect the results significantly since we added ML models in the pipeline to further identify relevant reviews. However, some inaccuracy is associated with models used to identify potential VPN-switching reviews and themes mentioned in any review. We have tried to reduce this error through hyperparameter tuning and model selection. We also established the correctness and suitability of our VPN-switching review dataset with additional manual analysis. We also acknowledge that ML models used for the identification of themes in reviews will not be able to identify new themes, but thematic saturation attained during thematic analysis suggests that there will be a very small number of such unidentified themes and potentially have a small impact on the validity of presented results. Furthermore, we did not conduct a survey—rather, we preferred collecting and analysing an ecologically valid dataset posted by users of their own volition. However, this limitation of data also indicates that our reported results are a lower bound on the identified reasons. Another limitation is posed by fake reviews on the App Store or Play Store, but as indicated in [48], the fraction of such reviews is too small to impact our results. Yet another limitation is that we will miss the users who only used desktop VPN apps.

We uncovered interesting and potentially generalizable reasons for VPN-switching using our ecologically valid dataset despite these limitations. We also acknowledge the lack of generalizability of results associated with review blogging websites, as we have only considered 6 popular websites, but since these are top recommendation websites used by users, the results are still relevant. In fact, our blog post analysis results might be impacted by SEO techniques employed by VPN recommendation websites [60], but we still used these blogs as a data source since a regular user looking for VPN recommendation will also be subject to these SEO techniques and hence, will see these same websites.

**Ethical considerations.** In this work, we collected data from the Google Play and Apple App Store and conducted a qualitative (thematic) analysis of these reviews. Since we analysed user-generated data from the wild, we tried our best to conduct our study ethically and protect users' privacy in our dataset. Our Institute does not have a formal IRB, so we carefully reviewed and followed best ethical practices while conducting this study. In line with earlier work, we noted that no registration was necessary to view App Store or Google Play store reviews [16], signifying it's an open forum. Our analyzed reviews did not divulge any personally identifiable

information. Also, as advised by earlier work to protect user privacy, we further hashed user names (thus anonymizing them) in our dataset to protect user privacy [10].

Next, we discuss related work regarding adopting and using VPN apps. Then, in section 3, we discuss our methodology involved in the collection and filtering of data, followed by details of thematic analysis (Section 4). Section 7 describes the results of our thematic analysis, and Section 8 compares desired features during VPN-switching and VPN adaptation. Section 9 describes the in-depth set of desirable features expressed by users during VPN-switching. Section 10 describes our analysis of VPN review websites. We discuss the implications in Section 12 and conclude in Section 13.

## 2 Related Work

We discuss related work on the user-centric view of VPNs in two broad directions.

**Prior research of VPN awareness and usage.** Virtual Private Network (VPN) has been studied rigorously by prior research because of their vital role as Privacy Enhancing Technology (PET) and for other purposes like censorship circumvention, bypassing geo-blocking, etc. Most studies have focused on people’s awareness about the existence/usage of VPN [5, 29, 33] and on security-related concerns associated with VPN like anonymization, DNS leaks, etc [7, 28, 31, 32, 46]. A recent study by Dutkowska-Zuk et al. [14] conducted a study of 349 university students. They found that the students were mainly motivated to use VPNs to access restricted content rather than privacy. These studies about awareness and usage of VPNs were conducted through interviews [5] and surveys [29] while later were studied either by developing a pipeline for testing VPN apps [46] or by installing the app on a target device and then running tests on it [58] (an inherently invasive approach). Our study is complementary to this prior research as we focus on existing VPN users and identify their desired VPN characteristics that resulted in VPN-switching. Furthermore, contrary to the prior studies, we did not leverage user interviews or surveys [29]. Rather, we leveraged app reviews of popular VPN apps from the Google Play and Apple App stores, where users organically present their opinions and attitudes without researcher intervention, enhancing the ecological validity of results.

**Prior research of VPN adoption.** On the user side, data leaks, geoblocking services and awareness of privacy risks gave way to the wide adoption of VPNs [40]. Namara et al. [43] conducted a study about the adoption and usage of VPN apps and the barriers users encountered towards their adoption. Their survey of 90 participants found users’ emotional reasons to adopt VPNs—like surveillance or desire for privacy. A closely related study by Sombatruang et al. [54] checked factors influencing users’ decision to adopt a particular VPN application. They found user review ratings and

pricing significantly affected the user’s decision to choose a specific VPN app. Their results align with ours, with financial incentives and user reviews also playing a role in users’ VPN-switching decisions. However, other factors like Geographic location and interaction of the VPN providers with users also played a significant role in our case. Overall, our work is complementary to these earlier studies. They investigated why users adopt VPNs. On the other hand, our study we investigated what factors motivate users of a current VPN app to switch to another VPN app. In fact, we demonstrate that these users seek more advanced and customized features than those uncovered in previous work. Our work also builds upon work by Ramesh et al. [50], which conducted end-user surveys and interviews with VPN service providers. They investigated users’ dependence on VPN review websites for selecting VPNs and how these websites are not good sources of information. In that line, we unearth underlying reasons for VPN-switching and investigate VPN service recommendation websites to find the misalignment of user requirements and content provided by these websites. Next, we present our semi-automated approach of creating a review dataset mentioning VPN-switching in the wild.

## 3 Collecting VPN App Reviews Corresponding to VPN Service Switching

Our study’s first step was collecting app reviews of the most popular VPN apps. To that end, we turn our attention towards the Google Play Store and Apple App Store, which host all available VPN mobile apps for the Android and iOS platforms, respectively and also contain a user review section where users post their perceptions and attitudes towards these apps. This section describes the stages associated with collecting and filtering reviews in detail.

### 3.1 Collecting reviews for popular VPN apps

**Selecting VPN apps.** We started with Google Play Store and we first searched for “VPN” in app stores. However, the search results repeatedly contained the same VPN providers (e.g., VPN app, VPN-enabled browser app, lite versions). On the contrary, searching for “best VPN 2023” gave the most popular versions of individual apps. So, we searched for “best VPN 2023” and noted the top 50 VPN apps (in terms of downloads) which appeared in the search results. To ensure the diversity of the user base of VPN apps, We bucketed these apps by their number of downloads into 5 buckets—50K to 100K, 100K to 500K, 500K to 1M, 1M to 5M and 5M to 100M or more. and in order to study the variation of VPN-switching with popularity, instead of checking only top VPNs, we did stratified sampling from popularity buckets and ended up with 20 VPNs; they are—Hotspot Shield, Nord, Express, Vypr, Freedome, Tunnelbear, Windscribe, HideMyAss, Surfshark,

Surfeasy, VPN Unlimited, Zenmate, Proton, Hola, Cyberghost, IP Vanish, Strong VPN, Private Internet Access (PIA), Super VPN, Turbo VPN. To further test the popularity of these VPN apps, we compiled a list of the most discussed VPNs from VPN market report statistics [42, 59]—our set of 20 VPNs have a 75% overlap with this list.

**Collecting reviews for the VPN apps.** We used open-source tools to collect public reviews written in English (for simplifying analysis) from the Google Play store and Apple app store [37, 47]. In total, we collected 1,333,160 unique reviews during our data collection. The selected VPN apps, the number of downloads on Google Play Store and the number of reviews (on both platforms) are in Table 7 (Appendix A).

### 3.2 Identifying reviews mentioning VPN-switching

First, we pre-processed the review dataset (to retain only meaningful ones) and then used a simple heuristic to identify reviews explicitly revealing VPN app-switching.

Table 1: Number of VPN app reviews identified after our preprocessing step and application of heuristic.

Number of VPNs used	20
#reviews	1,333,160
#reviews with #words > 5	411,751
#reviews containing multiple VPN names	4,403

**Pre-processing reviews.** Guzman et al. [21] shows that short reviews are often non-informative and contain praise or dispraise without reason. Thus, we removed all reviews with fewer than five words to retain only meaningful reviews. Finally, we ended up with 266,895 reviews from the Google Play Store and 144,856 reviews from the Apple App Store.

**Detecting reviews possibly mentioning VPN-switch.** Note that most of the 411,751 reviews are likely written about a particular VPN app. To that end, to uncover potential reviews, we used a simple idea: *the reviews that contain names of multiple VPNs might indicate that the user is comparing them.* Thus, they might have used or considered using both apps (signalling switching). We leverage this idea to filter the reviews via a simple keyword search—we extracted keywords from each VPN app name (Table 9 of Appendix B contains the full list of keywords) identified reviews with keywords from multiple VPN apps. We identified 4,403 such reviews. Table 1 summarizes the number of VPN app reviews at each step of our process.

**Manual analysis to create final review dataset mentioning VPN-switch.** Since mentioning multiple VPN app names might not be enough to actually mention VPN-switching, at this point, we resorted to manual analysis. We randomly sampled 1,600 reviews out of these 4,403, and one researcher manually reviewed each of these reviews to classify the re-

Table 2: #reviews stating Actual switch, potential switch and neither (irrelevant) after our manual analysis of reviews mentioning multiple VPN app names.

Review type	# Reviews	#words		#characters	
		Avg.	s.d	Avg.	s.d
Actual switch	454	59.6	45.8	318.7	206.1
Potential switch	851	47.8	38.2	257.9	209.1
Irrelevant	295	25.3	39.1	132.8	203.2

views into three categories—Actual switch, Potential Switch and Irrelevant. For *Actual switch* reviews, a user explicitly mentioned that they switched between VPNs. For example, P661 mentioned “*I left [VPN1] for these guys! [VPN1] was slow.. and i mean slow... tried [VPN2]...and never looked back. love the speed and security!!*”. P2496 also mentioned “*Not able to connect to campus from dubai... bye to incompetent customer team..*”. In *potential switch* reviews, users showed intent of switching by comparing two VPNs but did not confirm switching VPN providers (however, we consider that if these users switched, they would switch from the VPN, which they are talking negatively about in the review). This comparison can be there without mention of a second VPN; for example, P1459 mentioned “*This is so stupid vpn I’ll connect from Pakistan... unable to Netflix... frnds able to watch!! useless!!*”. Irrelevant reviews are the ones that neither compared multiple VPNs nor mentioned switching. Out of 1,600 reviews, 1,305 (81.3%) indicated either an actual or potential switch. To check the correctness of annotation, we randomly sampled 100 reviews out of 1600 reviews and had them re-annotated by another independent annotator, and found Cohen Kappa between these two annotations to be 0.74, indicating a substantial agreement. Table 2 summarizes the number of reviews in each category as well as the average number of words and characters for each of these types of reviews. Interestingly, both the actual and potential reviews have a considerably higher number of words and characters than the irrelevant reviews. This observation hints that users provide detailed and potentially better explanations in actual-switch and potential-switch reviews. Next, we perform a multi-stage qualitative analysis to identify the reasons for VPN-switching revealed in these 1,305 reviews.

## 4 Identifying Reasons of VPN-switching Using Qualitative Analysis

We leveraged our dataset of 1,305 VPN-switching reviews to investigate the reasons for VPN-switching via qualitative analysis. To that end, two researchers together extracted explanatory quotes from these 1,305 reviews—in total, there were 2,808 final quotes for our analysis, which at least one researcher deemed explanatory for VPN-switching. Next, we use open coding and affinity diagram [53] to develop a hierarchy of themes explaining reasons for VPN-switching.



Table 3: First three levels of our hierarchical themes explaining the reasons for switching VPNs.

A Network	B.5 Deciet	F.1.1 Payment medium
A.1 Network speed	B.5.1 Removal of app reviews	F.1.2 Subscription options
A.1.1 Network speed	B.5.2 Financial Fraud	F.2 Price of VPN
A.1.2 Ping value	B.5.3 False claims	F.3 Refund
A.2 Server count	B.5.4 User data usage	F.4 Premium version of app
A.3 Server switching	B.6 Reliability	F.5 Free version of app
A.3.1 Ease of switching server	B.7 Reputation	G Dynamic data services
A.3.2 Manual switching required	B.7.1 Branding	G.1 OTT
A.3.3 Issues with server selection	B.7.2 Online reviews	G.2 Games
A.3.4 Issues with country selection	B.7.3 History	G.3 Ads
A.4 Network Protocol	C Connection	G.4 Streaming
A.4.1 Port forwarding	C.1 Connection speed	H Software
A.4.2 Support for IPv6	C.2 Connection stability	H.1 Updates
A.4.3 Wireguard	C.3 Split tunneling	H.2 Features
A.4.4 Issues with TCP/UDP	C.4 Ease of reconnection	H.2.1 Number of features
B User Interaction	C.5 Auto-connect feature	H.2.2 Presence/Absence of features
B.1 Assistance	D Geography	H.3 Devices
B.1.1 Technical support	E Security and privacy	H.4 Bypassing capabilities
B.1.2 Customer support	E.1 Security	H.5 Resource consumption
B.2 Services Provided	E.1.1 Encyption standard	H.5.1 Battery consumption
B.2.1 Amount of data available	E.1.2 App security	H.5.2 System resources consumption
B.2.2 Support for multiple devices	E.1.3 Server security	H.6 Self sufficiency of software
B.2.3 Cross OS support	E.1.4 DNS leaks	H.7 Software stability
B.2.4 Compatibility with environment	E.1.5 Headquarter country of VPN	
B.2.5 Issues with third-party apps	E.1.6 User activity logs	
B.2.6 Support for Always-On-VPN	E.1.7 Kill switch	
B.3 UI/UX	E.2 Privacy	
B.3.1 Mobile app	E.2.1 RTC leaks	
B.3.2 General UI issues	E.2.2 Usage access Permission	
B.3.3 General UX issues	E.2.3 Privacy policy of VPN	
B.4 User friendliness	E.2.4 SSH issues	
B.4.1 Ease of setup	E.2.5 Obfuscated servers present	
B.4.2 Ease of use	F Financial Aspects	
B.4.3 Accessibility	F.1 Payment	

**Open coding.** First, we applied open coding on the quotes. We randomly sampled 200 quotes, and using these quotes, two researchers collaboratively created a codebook. Apart from this, we set aside 50 quotes initially to check the saturation of themes. Next, using this codebook, those two researchers independently coded all the reviews. Inter-rater agreement (Cohen’s kappa) at the end of the open coding round was 0.81, indicating almost perfect agreement. At the end of the open coding round, the two researchers met to discuss the specific user quotes whose codes were mismatched and assign a final code to each quote (if necessary, creating new codes). After completing coding all quotes (apart from the random 50 quotes we set aside earlier), one researcher verified that the identified themes are sufficient to code those 50 quotes, indicating thematic saturation of our open coding. Finally, the coders assigned a total of 217 codes to 1,305 user quotes.

**Affinity diagramming to identify the hierarchy of themes.** After this open coding round, the two researchers collaboratively analyzed the uncovered codes using *affinity diagramming* where they together looked into the set of quotes for each code, not only the code itself [22]. First, we set aside 10 random codes to check the saturation of hierarchy. Then, the coders collaboratively created a set of higher-level themes from the rest of the open codes. They continued this process on the newly developed higher-level themes for two more rounds when the coders felt no new higher-level theme could emerge. At the end of the process, we identified a 4-level hi-

erarchy where Level-1 contains abstract highest-level themes created in the third (final) round of affinity diagramming, and Level-4 contains the very specific codes specified by the open coding step. Finally, we checked that including the 10 random codes did not add any new Level-1 and Level-2 themes, indicating thematic saturation of affinity diagramming. Our hierarchy of themes (first three levels) explaining the themes for reasons of VPN-switching is shown in Table 3. Level 1 contains the eight highest-level themes—Network, User interaction, Connection, Geography, Security and Privacy, Financial aspects, Dynamic data services and Software. However, we note that, so far, our analysis primarily considered only 1,305 semi-automatically filtered reviews (the complete hierarchy of themes is in Table 10 of Appendix C). Now, we scale our VPN-switching review dataset with themes using automated classification.

## 5 Scaling VPN-switching Review Dataset and Theme Annotation

We note that our VPN-switching review dataset comprises only 1,305 reviews (0.3% of a total of 411,751 reviews). To that end, we hypothesize that we might be missing a large-number of VPN-switching reviews which do not mention names of multiple VPNs. One such review in our corpus is “[VPN1] had been working great for me in the past but after last

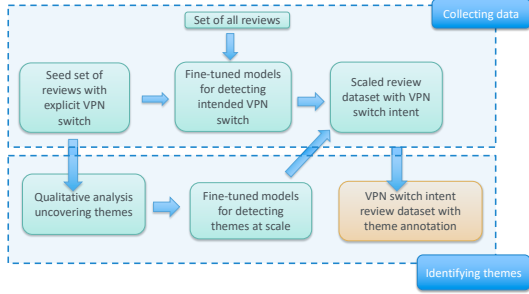


Figure 1: Pipeline for filtering review and identifying themes occurring in the reviews.

update it sucks...will look for other vpn if bug not fixed”. Thus, to identify such reviews, we designed a machine learning pipeline presented in figure 1. This section will describe the pipeline’s working and the resultant dataset.

## 5.1 Leveraging text-classification to extend VPN-switching review dataset

Originally, we manually divided 1,600 reviews into three categories—Actual switching, Potential switching and Irrelevant (Section 3.2). We note that we can consider these as three distinct classes assigned to the reviews—based on this idea, we model the problem of finding reviews mentioning VPN switching as a three-class classification problem. To that end, we tried multiple models as classifiers, including recent models like LLAMA2 [57], and FLAN-T5 [8]. The performance of individual models has been shown in Table 4. We used the fine-tuned DeBERTa model for further analysis since it offered best performance (F1-score of 0.85).

**Model Architecture.** Given the small number of labelled data points, we primarily used following pre-trained models, LLAMA-2 [57], FLAN-T5 [8], BERT [13], DeBERTa [23] and GAN-BERT [11]. We fine-tuned these model weights on the dataset for the downstream task of review classification. We used Optuna framework [26] for systematic hyperparameter tuning. Given the varying spectrum of models, the usage method also varied among the models (details in Appendix D).

**Model Performance.** We randomly split the dataset into 80% training and 20% validation subsets and post-training observed 85% accuracy and 84.73% F1-score on the validation dataset using DeBERTa model (presented in Table 4).

To further investigate the performance and misclassification of our model, we randomly selected an additional 100 reviews mentioning multiple VPN names (excluding the 1,600 annotated reviews), and one annotator manually annotated the selected reviews to create ground truth. Next, we used our text classification on these 100 reviews and constructed the confusion matrix (Table-5). We note that, in this confusion matrix,

Table 4: Performance of our fine-tuned models on the validation dataset for classifying reviews into Actual/Potential switch and irrelevant. Since the problem of review classification was a multi-class classification problem so metrics are computed using the methodology described in [19], where we compute model prediction using argmax among scores assigned by the model to all classes.

Model	Accuracy	Macro avg F1
FLAN T5	0.77	0.77
<b>DeBERTa</b>	<b>0.85</b>	<b>0.85</b>
LLAMA2	0.69	0.65
BERT	0.81	0.79
GANBERT	0.83	0.75

Table 5: Confusion matrix for our DeBERTa text-classifier annotation and ground truth for previously unseen 100 reviews. For only 9 reviews Actual/Potential switch is misclassified as irrelevant or vice-versa (marked in gray).

		Predicted		
		Actual switch	Potential switch	Irrelevant
Ground Truth	Actual switch	33	3	1
	Potential switch	13	20	2
	Irrelevant	2	4	22

in only 9 cases, a review mentioning an Actual or potential switch is misclassified as irrelevant. Thus, we identify actual or potential VPN-switching reviews in 91% of cases. Many misclassified reviews with strong negative/positive sentiments were not directed towards VPN features. Finally, we used this validated text-classified model on our dataset of 411,751 reviews to identify a total of 185,399 reviews mentioning actual or potential VPN-switching as presented in Table 6.

## 5.2 Identifying themes mentioned in extended dataset

Next, we wanted to assign themes to the reviews in our extended dataset of 185,399 reviews. To that end, we leveraged a similar idea as before—we modelled this problem as a theme identification problem and multi-label classification problem. Specifically, we model the Level-2 themes as labels assigned to each review. We chose to build our analysis based on L-2 themes—L-1 themes appeared too broad for granular analysis. Also, L-3 and L-4 themes might be more granular than L-2

Table 6: Final size of our extended dataset for VPN-switching reviews.

	Apple App Store reviews	Google Play Store reviews	Total reviews
<b>Actual switch</b>	38,139	52,166	90,305
<b>Potential switch</b>	35,614	59,480	95,094
<b>Irrelevant</b>	71,103	119,685	190,788

themes, but we have a relatively low amount of labelled data per L-3/L-4 theme. Thus automated theme identification with high accuracy would have been very difficult for them.

**Model Architecture.** We again used the following pre-trained models, LLAMA-2 [57], BART [35], FLAN-T5 [8], BERT [13], DeBERTa [23] and GAN-BERT [11]. Given the different architectures of these models, we fine-tuned BERT and GAN-BERT with multi-head classification layers on top of the pre-trained model. FLAN-T5 and BART were fine-tuned on the seq2seq task, and LLAMA2 was fine-tuned on the CausalLM task. We used Optuna framework [26] for systematic hyperparameter tuning (details in Appendix E).

**Model Performance.** To measure the performance of the fine-tuned models, we split the labelled dataset into 80% training and 20% validation subsets. We found that the BART model outperformed other models, identifying more than 80% of themes correctly in the reviews. The accuracy of all models on the validation set is shown in Table 20 of Appendix E. To further ensure the correctness of the model, we also performed a manual analysis—we selected 100 random reviews (excluding the original 1,305 annotated) and manually annotated them with Level-2 themes. Then, we compared the correctness of our predicted themes against this ground truth. The result of the manual analysis is detailed in Table 21 of Appendix E. Our classifier was correct for the majority of the reviews (more than 80%) in both the validation set and additional ground truth dataset. Next, we will leverage this extended dataset of theme-annotated reviews to explore our research questions.

## 6 Do Users Switch Between VPN Apps? (RQ1)

In our extended dataset, we found that 185,399 reviews contain user intent of VPN-switching out of 411,751 reviews, i.e., for a significant fraction of 45% of reviews. We further ask if unique users contribute these reviews and if the VPN-switching behaviour is uniform across different VPN apps.

**How many users are switching VPN apps?** We consider usernames as proxies for user accounts. Our collected dataset shows that the 185,399 VPN-switching reviews were posted by 169,201 unique usernames. The only exception is the “A google user” tag in place of the name, which represents user accounts that have either been deleted or anonymized due to the user’s privacy settings. Thus, these actual or potential VPN-switching reviews are posted mainly by individual user accounts (as opposed to a small subset of userbase).

**Are VPN apps uniformly affected by VPN-switching?** We further checked if the phenomena of VPN-switching is uniform across all VPN apps. To investigate, we checked what fraction of reviews per VPN app mention VPN-switching. Interestingly, there is a wide variation. Super VPN has only 2.0% of the reviews mentioned VPN-switching, and for VPN unlimited 61.9% of the reviews mentioned VPN-

switching—showing a 30 times increase. The percentage of VPN-switching reviews for each of our VPN apps is in Table 8 (Appendix A). To that end, we computed the Pearson product-moment correlation coefficient between the number of downloads for each VPN and the percentage of reviews mentioning VPN-switch. We found there is a statistically significant negative correlation ( $R = -0.5842$ ,  $p = 0.006858$ ) between these two variables. Thus, VPN apps with smaller user bases have a higher chance of their users switching to other apps. Next, we investigate why users switch VPNs.

## 7 Reasons of VPN-Switching (RQ2)

In this section, we delve into each of the eight level-1 themes (i.e., highest-level themes) to understand how the switching of users can be explained based on that theme. According to our study, the reason for the switch can be among a large number of causes ranging from Geography, i.e. the country in which the user resides or visits, to User interaction, i.e. how VPN service provider interacts with users through channels like customer service or through software/app. In this section, we will not use the names of VPNs in this section, as our goal is to understand the high-level reasons for users to switch VPNs, not to identify the specific shortcomings/superiority of any VPN. In this section, we will discuss the themes which have been identified in the manual analysis.

### 7.1 Security and privacy

VPN userbase comprises users who want to protect their privacy against hackers and government agencies. Thus, switching from one VPN to another seems evident for users who think their security is under threat. In our manual analysis, security and privacy appeared in 196 reviews out of 1,305 manually coded reviews. In this theme, we described particular aspects of security and privacy that were most dominant in motivating users to switch from one VPN to another.

**Hiding IP address:** Commercial VPNs typically function by tunnelling the user’s Internet traffic through a trusted remote server before it is forwarded to its final destination. This achieves the goal that the destination server does not learn the real IP address of the client [7]. Still, sometimes this fails to happen, and the actual IP address becomes visible to the destination server making the client vulnerable. This seemed to be a very relevant thing motivating users to switch from one VPN to another. For example, *P148* mentions, “Doesn’t work. Currently in Germany, and prime video knows I’m in Germany. After selecting US- New York or Washington DC, still knows I’m in Germany... [VPN1] had been the only one that works for me”, i.e. user feels their current VPN fails to hide their IP as Prime Video knows their correct location; thus, this user may switch to VPN1. We found quotes related to this theme in 40 reviews during our manual analysis.

**User logs:** User activity logs were another essential thing which users seemed to be concerned about because it was observed by users that VPN stored logs even if VPN service providers claimed that no logs were kept. In our manual analysis, user activity logs-related quotes appeared in 19 reviews. Also, many users expressed concerns about cases of selling logs, which threatened their privacy. One such example can be seen when *P10 mentioned*, “*It (VPN1) was caught selling user’s data and logs in 2017. Use at ur own risk. Use [VPN2] instead of this*”, clearly warning users against VPN1 as it has a history involving the sale of user logs and data.

**Headquarter country of VPN:** One surprising yet recurring reason (found in 20 reviews during the manual analysis motivating users to switch from one VPN to another was the country where the VPN is headquartered. Many users felt that companies with headquarters in some countries were more likely to betray them than other companies. For example, *P899 mentioned*, “*...I let it run out. Why? Because [VPN1] is AMERICAN, and I’m done trusting Americans with my data...I’ve switched to [VPN2], a Swiss company. So far so good*”, where the user switched from VPN1 to VPN2 just because VPN1 is an American country while VPN2 is not. In our corpus, users were often against companies headquartered in “14 eyes countries”.

**App usage access:** Permission control is one of the major Android privacy mechanisms. When an application is to be installed, a user has a choice whether to allow specific permissions or not [18]. While getting installed, many apps ask for “Usage Access” permission, allowing the app to track what other apps you’re using and how often, as well as your carrier, language settings, and other details [56]. This appeared as a concern to some users. We found quotes related to this theme in 6 reviews. For example, *P30 mentioned*, “*Why does it need “App Usage Access” permission? [VPN1] doesnt ask for it...*”.

**DNS leaks:** The Domain Name System (DNS) is an essential Internet service that translates domain names into IP addresses [41]. Information leakage of DNS servers may cause the disclosure of users’ behaviour or network zone structure and may even induce phishing or intranet penetration attacks [38]. DNS leaks were also raised as concerns in reviews mentioning them as reasons for switching from VPN. Other fundamental issues raised by users were related to issues with SSH, webRTC leak, etc. During our manual analysis, we found this concern in 10 reviews.

## 7.2 User demographics

Now we check how user demographics, such as place of residence, financial constraints, etc., are also responsible for making users switch between VPNs.

### 7.2.1 Geographic location of user

Due to the increase in awareness about the usage of VPNs as a Privacy Enhancing Technique and the rise in geoblocking imposed by various governments, there has been a worldwide increase in usage of VPNs, but this is sometimes constrained due to poor performance of VPNs in certain countries. In our study, we found this lack of performance in certain locations to be one of the reasons why users switch from one VPN to another (expressed in 88 out of 1,305 manually-coded reviews). For example, *P818 mentioned*, “*[VPN1] hasn’t worked in China for a year. The app is useless on the phone and computer. The back up open listings are inadequate... Will never subscribe again. [VPN2]’s free service has worked better*”. Apart from the lack of customer care service, the failure of [VPN1] to work properly in China was an important reason why users were willing to switch to [VPN2].

**User residing in China:** Among the censorship regimes on the Internet, China is one of the most notorious, having developed an advanced filtering system known as the Great Firewall (GFW) to control the the flow of online information [24]. The strength of this firewall was also reflected in our study, as most reviews related to this theme were related to China (50 out of 88 Geography-related reviews), followed by Middle East countries like Saudi Arabia and UAE. In China, users complained of VPNs not upgrading their services as their servers were blocked as reflected in *review of P835 mentioning* “*anywhere else, hong kong for instance, connect right away. Now in nanchang, china, seldom succeed, because censors keep attacking and cutting access. [VPN1] stays ahead by highly frequent updates. perhaps [VPN2] could do something similar*”, where the user is currently using [VPN2] and expects [VPN2] to introduce updates frequently, like [VPN1], to fight censorship.

**User residing in Middle East countries:** Similar patterns were observed in Middle East countries, according to [45], blocked content often contains messaging that opposes the censoring government’s own narrative of the conflicts, and that governments block content that could potentially spark internal dissent. We found quotes mentioning these countries in 12 reviews under this theme. Thus, VPN users in these regions face the unavailability of proper services and switch from one VPN to another. For example, *P139 mentioned in their review*, “*Cannot get the app to connect from UAE to USA. The chrome extensions says unable to load. Tried [VPN1] and that one worked. It even made [VPN1] run. Looks like [VPN1] is blocked in UAE. Otherwise it worked great in Europe*”.

**Other locations:** Other than these two regions, reviews also mentioned problems faced by users in Australia, Europe and South East Asia. One reason was the inactivity of VPN after government intervention, e.g., in Turkey the government blocked some servers as expressed by *P529 mentioning* “*Works Flawlessly. I’m in Turkey and the government blocked [VPN1] so I tried [VPN2]. It is fast and always works. Highly*



recommend”. Within this theme, we noted that this inaction of VPNs towards government policies (e.g., blocking specific services) was the main motivation to switch VPNs.

### 7.2.2 Financial constraints of user

Most VPN services worldwide provide either premium or free service, or both leaving the choice to the user about the type of service they want to procure. Companies that provide users with the option of both free and premium versions of their VPN incentivize users to buy the premium version by increasing the daily usage limit or the number of simultaneous connections users can have. Some VPNs also come with a free trial period where users can try the premium features for a limited period (typically 7 days) and then evaluate if the user’s needs are satisfied by this VPN. The sub-themes for this theme cover the contribution of these aspects of VPN service in motivating the user to switch. In our manual analysis, we found quotes related to this theme in 294 reviews out of 1,305 manually-coded reviews. Specifically, users mentioned the following reasons.

**Pricing of VPN:** High pricing of VPN was a recurring theme for some VPNs and sometimes served as the primary reason for users switching to another VPN. During our manual analysis, we found quotes related to this theme in 182 reviews. *P559 mentioned, “Good service but just too expensive. I have switched to [VPN1] and saved a fortune. I’m afraid that with such capable competitors offering such good value, [VPN2] will have to tidy up their pricing”*. Here, it is clear that the user switched from VPN2 to VPN1 even though the user feels that the service provided by VPN2 is good. While the pricing of VPN was a big issue for some users, a proportion of users cared more about the service provided than the price, for example, in *“This is probably the best VPN on the market. It costs more than other ‘budget’ VPNs but you get what you pay for. The speeds are phenomenal, especially for someone like me coming from [VPN] where I was paying to sit and look at the buffering symbol on YouTube...”*, where the user is ready to pay more and believes the service justifies the cost.

**Free trial:** Companies often provide free trial to lure users towards buying a premium version of VPN. However, it sometimes backfires—*review of P1478 mentioned “I tested this as an alternative to [VPN1] when my subscription ended, however they did not give the full 7 day trial, which instead ended after 5 days. The performance was okay, but that trial bug made trust them less and in the end I went back to [VPN1]”*, here the bug in the free trial was disappointing for the user; thus, the user moved back to the previous VPN. Some companies ask users to enter payment information or make some initial payment for accessing a free trial but it also has some repercussions, as in, *review of P145 mentioning “Expensive and the Free Trial isnt a free trial. It still makes you put in billing information. That’s just wrong. Never try this. They shouldnt have your info until you’re ready to commit. That’s*

*not a free trial idc if it allows you 7 days to cancel. Most people forget and thats what they’re counting on”*. Here, we can see that the user is just switching due to distrust that VPN3 might be a fraud, and that’s why they are asking for payment, which isn’t what these payments are intended for. We found quotes related to this theme in 41 reviews during our analysis.

**User finding no perk of paying for premium:** One interesting theme observed from reviews is some users, after buying premium features, felt no difference from the free version of the VPN. During our manual analysis, we found quotes related to this theme in 10 reviews. *P745 mentioned, “from the first day I buy a one month premium subscription I felt like it was the same as the free version of the app. now I regret buying this piece of junk and thinking that I have to deal with it for the whole month. makes me angry and I feel like a stupid not going for the [VPN1]”*. Thus some VPN services are either not able to convey all features of the premium version to the users or it is not providing anticipated features to users.

**General issues related to finances:** Problems with a payment portal, lack of subscription options, lack of offers for existing users, etc. were other factors noted in our study as visible in reviews, for example, *P340 mentioned, “The description states that it has in app purchases but there is no way to pay from your Google play account. I had been saving money in my account for this purchase and now I find out I can’t use it for this app. if you want to use your Google Play balance look into “[VPN1]” they will let you pay with your Google play balance”*. Apart from these, some positive factors also motivate users to switch from their current VPN to that particular VPN, for example, the availability of more subscription options or lower price as in their review, *P258 mentioned “Love it. Cheaper than [VPN], more subscription options than [VPN]. So easy to use”*. During our manual analysis, we found quotes related to this theme in 57 reviews.

## 7.3 Software usability

While the primary purpose of VPNs is to protect the user from privacy attacks while maintaining anonymity but usability is also an important factor considered by users. Our study found that many users switched among VPNs due to usability issues like poor UI, lack of required features, etc. In our manual analysis, we found quotes related to this theme in 524 reviews out of 1,305 reviews. Out of many recurring reasons, we have described a few of the most frequently appearing reasons responsible for users switching between VPNs.

**UI/UX:** User experience covered under UI/UX also appears as an essential aspect of this theme. Better UI serves as an important reason motivating users to switch between VPNs. During our manual analysis, we found quotes related to this theme in 129 reviews. Also, we observed that many users admired the UI features provided by other VPNs. For example, *P360 said, “Pretty basic app but it works well. I would like*

a quick toggle in the notification to pull down like [VPN1]”, where the user admires the feature provided by VPN1. In some cases, we also observed that this admiration is so extreme that users tend to switch just because of them. Bugs in the user interface also seemed to matter for users motivating them to switch if not solved. For example, P491 mentioned, “I just installed your VPN on my phone. However, I am uninstalling it now. There is no way to go back to the home screen once you connect to [VPN1]. If you go back, you get disconnected. This is disgusting. Take some cue from [VPN2] where there is an on or off switch...”, where this bug in UI makes the VPN useless for the user and hence the switching.

**Presence of too many features:** While the lack of some basic features bothered users, the presence of too many features was also mentioned as the reason for switching between VPNs by some users, as with the number of features available in a software increase, resource consumption of software is also generally observed to increase. For example, P54 said, “So sick of this VPN, they keep adding intrusive and unwanted features. I never wanted a file cleaner or task killer in my VPN app. Uninstalling and switching to [VPN2]”, here the user is switching to VPN2 primarily because the current VPN has a lot of overhead in terms of features like file cleaner, task cleaner, etc. A similar pattern was observed in User Interface, where we found both very simple and very complex UI motivating users to switch VPNs. We found quotes related to this theme in 17 reviews.

**Connection speed:** We found speed is a significant factor for users switching between VPNs. While network speed is related to the latency experienced by users when the user is connected with a VPN, the connection speed is the time the VPN takes to connect to the desired service (e.g., a webpage). Some users mentioned slow connection speed as a primary motivation for VPN-switching. For example, P769 mentioned, “You better take this review because I have used [VPN1] for three years with friends who use the same...[VPN1] is the worst at connecting & also log in...Connecting to country take hours. Other vpn connects in 5 sec. Just downgrade ur app, man” Clearly, the user mentioned [VPN1] has issues with logging in, and the connection speed was not up to the mark. During our manual analysis, we found quotes pertaining to this theme in 39 reviews.

**Network speed:** Network speed has emerged as the most commonly cited metric for characterizing the quality of network offerings [3]. Numerically speaking, this was the most common theme that arose from our corpus study. For example, P103 mentioned in their review “Nice app. The one thing that this app(VPN1) is lacking is the speed of the vpn. I’ve used [VPN2] for quite a while but this speed is pretty slow...”—the user was directly comparing speeds provided by VPN1 and VPN2. During our manual analysis, we found quotes pertaining to this theme in 375 reviews.

**Software updates:** Software updates are essential to keeping

systems and programs up-to-date. These updates fix bugs and improve performance and usability, but arguably their most important function is enhancing system security by fixing vulnerabilities [39]. In our study, we observed that infrequent updates are criticised by users as P963 mentioned in their review “...It’s been unreliable for weeks and the last mobile update was more than a fortnight ago. During the China blockages, [VPN2] had version updates and information for customers on their website”. Here the user felt VPN2 is superior because of their frequent updates.

Another important aspect arising in our study related to updates is when users think that updates introduced in apps make service worse for them. For example, in P1 says, “There are better VPN Apps. [VPN1] become dead. Try [VPN2]...[VPN1] they still old school and now these frequent updates, its get worse”, where users feel that the updates introduced in the app are making their user experience even worse. Also, we found certain reviews where, according to users, no persisting problems were solved by updates—P227 mentioned, “Users in China, don’t bother. I’ve had [VPN1] for 2 years, no idea why I renewed my subscription. You can’t connect more than half of the time. Constant updates that never seem to fix anything...”, where constant updates don’t seem to solve the network issues faced in China by the user. During our manual analysis, we found quotes related to this theme in 54 reviews.

## 7.4 Other frequently appearing factors motivating users to switch

In this sub-section, we will discuss other factors that frequently pop up in many user reviews. These themes are mostly related to the specific purposes for which users use VPNs, like accessing OTT platforms from restricted regions. The frequent appearance of these factors showed the extent to which VPN technology has been democratized.

**OTT Platforms:** The OTT service providers deliver audio, video and other media over the internet and bypass the traditional operator’s network. Since the OTT players do not require any business or technology affiliations with network operators to provide such services, they are often known by the term ‘Over-The-Top’ (OTT). Some popular OTT platforms were mentioned quite frequently in our study—Netflix, Amazon Prime Video, Hulu and Hotstar [2, 25, 27, 44]. During our manual analysis, we found quotes pertaining to this theme in 51 reviews.

Many users whose primary purpose was accessing the content of another country on these OTT platforms, which are generally blocked due to policies of government or OTT service providers, switch between VPNs if their current VPN does not provide expected access. One such example is seen when P1400 mentions in their review, “There is no streaming Netflix server for another country than the US as US don’t necessary have all title. [VPN1] has that..and able to con-

nect to any region might consider jumping back to [VPN1]”, where the user is considering switching from the current VPN to VPN1 because of restricted access to Netflix.

**Gaming issues:** By encrypting the internet connection and hiding the IP address, using a VPN for gaming can potentially help in several ways — granting access to different servers, improving ping and reducing lag, or even protecting you against DDoS attacks [34]. Thus, many gamers nowadays look forward to VPNs for a better and more secure gaming experience. Our study also found that many users switch between VPNs to get the expected gaming experience. For example, *P1395 mentioned in their review, “I was happy with this service until I realize that this service doesn’t work with online games. The games keep disconnecting. I thought it was game issues until I switch to [VPN1]. taa daa, no more online games disconnections...”*—the user switched VPNs because of disconnections in online games. During our manual analysis, we found quotes related to this theme in 8 reviews.

**Streaming issues:** Under this theme (included in 34 reviews), we also studied the importance of ease of streaming using VPNs in motivating users to switch from one VPN to another. Many streamers use a VPN to protect their actual IP address for privacy and performance reasons. Some use it to access Twitch or unblock content and games [12]. From the collected reviews, we found that many users considered switching to another VPN just because their current VPN doesn’t perform well with streaming. For example, *P126 mentioned, “...I found it unreliable for video streaming. It was extremely slow and videos often just stopped playing and I couldn’t get them to restart. I’ve switched to [VPN1]...”*, where the user switched to VPN1 because of issues with streaming videos.

**Technical and customer support offered:** Assistance provided by the company through technical support and customer support are the most common ways through which a company interacts with its current and potential customers. In our study, we discovered that poor knowledge of technical support, which made it difficult for users to solve the problems they faced while using VPNs, was an important reason for motivating users to switch. During our manual analysis, we found quotes related to this theme in 151 reviews. As *P146 mentioned, “I used [VPN1] for around 7 years. However, during the middle of 2020 I was unable to view TV anymore. After five months and numerous emails and promises that technicians were working on the issue, I gave up! I had read positive reviews of [VPN2] and decided to give it a go...”*, the inefficiency of technical support was a reason why the user switched from VPN1 to VPN2.

Next, we will discuss how these factors differ from those identified in the previously published literature on factors influencing the choice/usage of VPNs. We will analyze the impact of these factors on users’ decisions to choose VPNs while switching.

## 8 Comparison of Features influencing VPN-adaption and VPN-switching (RQ3)

Our study identified reasons for VPN-switching, however, it is not obvious if these reasons for VPN-switching are different from the general VPN adoption motivation of users. Especially since Sombatruang et al. examined factors which users consider while choosing a VPN app which suits their purpose and preferences [54], to that end, we contrasted those factors from earlier work with the ones identified in our study to find how user requirements and concerns regarding the choice of VPN change with prior usage of VPNs.

Sombatruang et al. found that factors like device battery consumption, change in internet speed due to VPN, connection stability, ease of connection-establishment of connection, etc., are primary considerations of users while choosing a VPN [54]. They also highlighted the importance of app store reviews and star ratings of apps in the user app selection procedure. The subscription price and User personality are also crucial in the VPN adaption process [43]. These themes are also present in reviews exhibiting intent to switch, but there are some key differences, possibly due to users’ experience.

**Advanced user requirements for VPN:** Unlike prior work, in our study, one recurring theme was that users complained about the lack of advanced features like split tunnelling, kill switch, better UI, streaming support while playing specific games, etc. Out of these, kill switch and split tunnelling security features are more common—we can infer that existing users understand the importance of these features after facing associated problems in their existing VPN. This informed understanding strongly influenced the VPN selection criteria during switching.

**Dependence on user Demographics:** Prior work found that the purpose of using a VPN is an important factor in the VPN selection process. We found that while this was still true, user demographics (specifically the geographical location of users) also played an important role in VPN switching. This was particularly visible in reviews from users residing in China and the Middle East, where web censorship is very high and VPNs are often blocked. In those situations, users are forced to switch between VPNs frequently if the app prevents them from evading geographical constraints.

**Reduced weightage to ratings and reviews:** Sombatruang et al. identified that ratings and reviews are most significantly correlated with VPN adaption, indicating herding attitude among users [54]. In contrast, our dataset had only 4.15% of reviews in which users talked about app ratings and user reviews of the VPN app, while other factors like geography, security, etc., are mentioned in significantly more reviews. Thus, our results hint that user experience with existing VPN apps often reduces the weightage of “ratings and reviews” as a factor during VPN-switching. Next, we will check the key features desired by users during VPN-switching.

## 9 VPN Features Desired By Users During VPN-switching (RQ4)

In this section, we check what features are desired by existing users during VPN-switching and if users desire multiple features together while switching their VPN providers.

**Top desirable features in a VPN.** We use our identified level-2 themes as features desired by users while switching their VPNs. To that end, we simply searched for these themes in the reviews and took the frequency of occurrence as importance for these features. The top-5 most mentioned features in VPN-switching reviews are Connection, Network speed, User-friendliness, Service Provided, and UI/UX appearing in 41,020, 28,406, 22,402, 21,689, 19,406 reviews, respectively (top-10 of such features with their frequency is in Table 25 of Appendix F). We noted that privacy is not among the top 10 mentioned features in VPN-switching reviews. This finding aligns with earlier work [54]. Next, we moved beyond only the most popular features and checked if these popular features also co-occurred with not-so-popular ones for specific VPN-switching reviews.

**Co-occurrence of desired features in VPN-switching reviews.** For checking the co-occurrence of multiple desired factors, in line with prior work, which examined the co-occurrence of tokens in the same posts [52], we created a co-occurrence network [36] of level-2 themes where themes served as nodes of the network and weights of edges were calculated using the formula:  $W_{ij} = F_{ij} / (F_i * F_j)$  where  $F_{ij}$  are #reviews in which both theme  $i$  and  $j$  occur,  $F_i$  is the number of reviews in which theme  $i$  occurred, and  $F_j$  are #reviews in which theme  $j$  occurred. After constructing the co-occurrence network, we used the clique-based Louvain algorithm [4, 20] to find 10 overlapping communities in the graph.

**Specific subsets of desired features co-occurred in VPN-switching reviews.** In our detected communities, we found that the most frequent factors in user reviews about switching VPN and network speed co-occurred with non-frequent factors like streaming. This shows that user requirements often encompass multiple features, and furthermore, there are distinct feature sets that are desired together in VPN-switching reviews, signifying a concrete spectrum of needs which led to VPN-switching. One interesting observation is that 7 out of 10 communities have “Security” (3 communities) and “Privacy” (5 communities) in them even though privacy did not occur most frequently in reviews—signifying at least a non-trivial subset of users still consider better security and privacy as a valid reason for switching VPNs. We provide the complete list of the feature co-occurrence communities for interested readers in Table 24 (Appendix F). Next, we finally check if the current VPN review blogs identify and cater to these distinct user requirements.

## 10 Alignment of Information in VPN Review Websites with User Requirement During VPN-switching (RQ5)

We previously noted in Section 8 that users switching among VPNs put lesser emphasis on reviews and ratings while switching. However, as previous work found, general users still rely on VPN review blog websites for information regarding VPNs [50]. To that end, we ask if these VPN-review blog articles can help the users who want to switch VPNs.

### 10.1 Collecting data from VPN-review blogs

In order to identify popular VPN-review websites and their relevant blog articles, first, we queried Google with the term “best VPN” and collected the top 10 results. Four out of these ten websites prohibit automatically collecting their articles, so we focus on the rest. We searched each website domain using Google’s site search with the keywords “Best VPN recommendation blogs”, “VPN review blogs”, and “VPN comparison blogs”, which resulted in 583 blog articles. We manually reviewed these articles to identify 376 blog articles related to VPN or its uses. Out of these, 176 are VPN review blogs, and 200 are VPN best-pick blogs (i.e. blogs which either recommend a ranked list of best VPNs or compare VPNs). For ethical reporting of results, this section will refer to individual sites with alphabets A to F.

### 10.2 Topic analysis of VPN-review blogs

**Topic Modelling:** To identify the content available in the blogs, we started with Topic Modelling implemented using LDA algorithm [30]. Firstly, we applied LDA to the text of the entire 376 blog corpus, followed by applying it to the corpus formed by only the VPN review blogs and VPN best-pick blogs separately.

**Identifying themes in reviews:** Apart from topic modelling, we also applied the BART model fine-tuned for identifying themes in reviews on the blogs corpus. Given the input length constraint of the BART model, we fed the blogs to the model in small chunks followed by a union of identified themes. This provided us with the themes covered by blogs from these VPN recommendation websites.

**Blogs do not give information on important and desired VPN features:** We started by applying the LDA topic modelling algorithm on our blog dataset corpus [30]. LDA output word clusters as topics. Following the method described in prior work, we identify the optimal number of topics and assigned semantic labels to these word clusters via manually checking the blogs where words from a cluster appear [6]. Finally, we identified ten topics in the blogs, out of which four topics were related to VPN features. Two topics primarily contained terms related to streaming services like “Hulu”,



“Netflix”, etc. The other two topics were related to “Security and Privacy” and “Connection and Network speed”. These are in line with the themes identified in our qualitative analysis.

We also noted a lack of coverage of other important aspects mentioned by users in reviews like financial aspects (like price, free version availability, etc.) and Geography. We also applied topic modelling on subsets of blog articles—VPN review blogs and VPN best-pick blogs separately. Topic analysis of VPN reviews and best-pick blogs yielded 10 topics each. Most of the topics in VPN-review blogs were related to Streaming and Speed, while topics in VPN best-pick blogs were mainly related to the Speed and Pricing of VPNs. However, these topics did not touch upon a plethora of features mentioned in Section 9—these blogs might provide information about a significant number of desired features which users consider during VPN-switching. The topics (semantic labels with word cluster) are in Table 28, 29 and 30 of Appendix F.

Then, we identified themes covered in the blogs using the BART model fine-tuned to identify themes in reviews to understand the coverage of user requirements by these blogs. The top themes identified by the model are listed in Table 27. Among the top themes, we can see “Privacy” and “Features”, which, according to our findings, are also the top aspects users look at when making the decision to switch VPNs, but we see no coverage of important themes like “Geography” or “User Friendliness” or “Reliability” which is also important for users. We also checked the variation of topics covered in VPN review blogs and VPN best-pick blogs and saw a similar trend as before. Best-pick blogs mostly focused on Speed and Pricing, and VPN review blogs focused on features but lacked coverage of other important aspects related to VPN switching.

### 10.3 Entities mentioned in VPN-review blogs

We Performed an off-the-shelf Named Entity Recognition technique (NER) [51] on individual blogs and then calculated the frequency of these entities in blogs from different websites as well as in the whole blog corpus. The full list of the top 10 most frequently occurring entities for each of the websites is presented in Table 22 of Appendix H.

**Information in blogs is heavily biased towards specific VPN apps and services:** Our frequency analysis of entities reveals that there are 1 to 5 specific VPN app names in the top 10 most occurring entities across websites. Out of these two entities occur very frequently—“Netflix” and “Surfshark” (Table 22 of Appendix H). We make two observations from this analysis. First, these websites focus disproportionately more on streaming services like Netflix, backing our observation in Topic modelling. Second, these websites primarily mention a few VPNs like Nord or SurfShark while ignoring other VPNs. However, interestingly, a Pearson correlation coefficient analysis did not find any statistically significant correlation between the frequency of the entities (which are

VPN apps) and the number of downloads. Finally, we do a word association analysis to check (i) if these mentions of specific entities in blogs (i.e., co-occurred) are biased (e.g., overwhelming praise or criticism) and (ii) if the themes identified in Section 9 (which are also user-desired features) are at all mentioned in any form in these blogs.

## 11 Bias towards specific entities and coverage of features in VPN-review blogs

**Word association analysis:** We leverage a technique developed in prior work to perform a word association analysis, i.e., given a set of words, what are similar or associated words appear in a corpus? [55]. We first created two sets of keywords. The first set contains keywords that are not related to any specific themes, and rather, they contain general terms describing desired VPN characteristics like “best”, “fastest” and “Cheapest”. The second set contained a set of keywords which are related to themes (i.e., features) desired in VPN-switching reviews. Next, to capture words used in similar contexts, we used a pre-trained BERT model [13] and fine-tuned it using the blog corpus on a Masked Language Modelling task to get a model acquainted with the context used in blogs. Then, we generated embeddings for keywords using this BERT model. Finally, we analysed embedding of which words are similar (as described in prior work [55]) to embeddings of our sets of keywords. This approach helped us to capture the context of input keywords within the blogs while also taking the general meaning of words into context due to the nature of the pre-trained model. All keywords are in Table 26 of Appendix H.

**Bias towards specific VPN apps in blogs:** For the keyword “best”, in most websites, the names of popular VPNs like Nord and Proton are among the most similar words. We made similar observations when we tried to find similar words for terms like “fastest” or “cheapest”, indicating a bias towards selected VPNs corresponding to all words in our first set of keywords—signifying desired VPN characteristics are always associated with these VPN apps in the blogs.

**Lack of coverage for user-desired features in VPN-blogs:** Leveraging our second set of theme-related keywords, we checked if VPN blogs used keywords related to user-desired themes (in VPN-switching reviews). Specifically, for each keyword in our set (e.g., “logs”, “free”, or “price”), we checked top 10 most similar words in the corpus (in terms of our fine-tuned BERT model embedding). However, for most themes, these words are not relevant—possibly because most of these themes are not discussed in depth in the VPN blogs.

## 12 Implications

Our first-ever study of VPN-switching study unearths a number of important observations. To that end, in this section,

we examine the implications of our study for VPN providers, VPN-review blogs and the security community in general.

**VPN-switching is intended by a non-trivial fraction of users:** We note that VPN-switching, although not studied before, is a significant phenomenon. More specifically, out of 411,751 reviews from 20 popular VPN apps, 185,399 (45%) indicated towards VPN-switching. This observation heavily emphasizes the need to assist these users from the VPN-review blogs as well as from VPN providers by providing them services that cater to these users' desired preferences.

**Users desire advanced and more customized features while switching VPN:** As opposed to VPN adoption, the users who want to switch VPNs are more knowledgeable. Thus, they seek more advanced security features and customized options (e.g., for evading the censorship employed in specific countries). Thus, VPN providers might want to provide customized solutions to these experienced users.

**Users desire sets of features while switching VPN:** Our cluster analysis revealed that, interestingly, the VPN-switching reviews often desire multiple features for their new VPN app. In fact, these features form distinct clusters, hinting at separate user bases (since mostly unique users posted these reviews) with distinct requirements. In fact, often during VPN-switching, security and privacy is an important factor, but it is accompanied by additional requirements in a VPN. This finding provides a roadmap for VPN providers and researchers to create customized VPN solutions which can cater to these distinct user bases with different sets of requirements.

**Current VPN-review blogs are insufficient for VPN-switching for bias toward specific VPNs and low coverage of user-desired features:** Finally, our analysis reveals a huge gap and opportunity for researchers and the general security education community. Currently, VPN review blogs paint a biased picture of VPN apps and manipulate users to switch to specific VPNs. In fact, presumably, due to this reason, the users who want to switch VPNs did not mention these reviews and ratings as significant factors for choosing a new VPN. Furthermore, these blogs are insufficient to provide information to users regarding their desired privacy, security and usability features. Hence, our findings imply that the security community should develop more tools like Vpnalyzer [49] and make them user-friendly so that users can check the VPNs and switch to a VPN according to their requirements.

## 12.1 Recommendations for stakeholders

Finally, we synthesize a list of recommendations for important stakeholders in the context of VPN switching—VPN providers, VPN recommendation websites, the Security Research Community, and finally, the users themselves.

**Recommendations for VPN service providers:** Our findings indicate that there is no one-size-fits-all thing in the case of VPN. Hence, service providers need to provide users with

options to customize VPN features based on users' priorities and requirements. The providers can also aid users by customizing their products according to user geography and demographic groups. In our study, we found that there is a large section of VPN users who expressed "Potential Shift" in their reviews. These users are basically giving their existing VPNs a chance to satisfy their need. Thus sensing these needs and establishing a smoother communication channel between user and providers (e.g., via customer or technical support) can help the company retain these users.

**Recommendations for VPN recommendation websites:** We identify that VPN recommendation websites are biased towards particular VPN service providers and do not cover important themes that users consider while switching VPNs. Therefore, coverage of such topics in blogs can help users make more informed decisions. We also noted a difference in factors responsible for VPN adoption and switching, which can guide these websites to tailor their blogs according to users' VPN usage experience. These websites can also provide automated VPN recommendation systems according to users' experience and requirements.

**Recommendations for the security research community:** Our results show that users sometimes compromise their security for their requirements, like watching Netflix. Thus the security community might consider developing more tools like Vpnalyzer [49] and make them usable so that users can test the VPNs by themselves.

**Recommendations for the users:** Finally, from previous literature, we found that users rely a lot on VPN recommendation websites [50], which is a problem. In our study, we found that these websites are biased towards and selected VPNs and hence, users should rely on more reliable sources like academic tools or materials to select the VPN. Also, users need to acknowledge that their peers or even families will have different requirements compared to theirs; hence, just relying on word-of-mouth might not work while selecting VPNs.

## 13 Conclusion

This study delved into the VPN-switching behaviours of users through an analysis of 185,399 app reviews from 20 popular VPN apps across Google and Apple stores. Our findings revealed a significant number of users intending to switch between VPNs, with varied rates among providers. Notably, the size of a VPN provider's user base exhibited a statistically significant negative correlation with the frequency of VPN-switching mentions in reviews. Our qualitative analysis identified a hierarchy of reasons motivating users to switch, encompassing technical factors such as network speed and non-technical considerations like financial aspects.

Furthermore, our research introduced a novel exploration of VPN review blogs, exposing biases towards certain VPNs and a notable absence of information regarding user requirements. By shedding light on the reasons behind user retention

and loss for VPN providers and unveiling patterns in user preferences, this study offers valuable insights for enhancing VPN services and addressing user needs more effectively.

## 14 Acknowledgement

We thank the anonymous reviewers and our shepherd for their valuable feedback. We also thank Sankalp Ramesh for his help with an earlier iteration of this work. This research was (partially) funded by a Google India Faculty Research Award.

## References

- [1] Susan Alexandra. 5 popular reasons for using a vpn. <https://www.globalsign.com/en-in/blog/5-popular-reasons-using-vpn>, 2020.
- [2] Amazon. Amazon prime video. <https://www.primevideo.com/>, 2023.
- [3] Steven Bauer, David Clark, and William Lehr. Understanding broadband speed measurements. 08 2010. [https://groups.csail.mit.edu/ana/Publications/Understanding\\_broadband\\_speed\\_measurements\\_bauer\\_clark\\_lehr\\_TPRC\\_2010.pdf](https://groups.csail.mit.edu/ana/Publications/Understanding_broadband_speed_measurements_bauer_clark_lehr_TPRC_2010.pdf).
- [4] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), Oct 2008. <https://arxiv.org/abs/0803.0476>.
- [5] Bram Bonné, Gustavo Ruiz, Peter Quax, and Wim Lamotte. Insecure network, unknown connection: Understanding wi-fi privacy assumptions of mobile device users. *Information (Switzerland)*, 8, 07 2017. <https://www.mdpi.com/2078-2489/8/3/76/htm>.
- [6] Ceren Budak. What happened? the spread of fake news publisher content during the 2016 u.s. presidential election. In *Proceedings of the Web Conference 2019*, WWW '19, page 139–150, 2021.
- [7] Thanh Bui, Siddharth Prakash Rao, Markku Antikainen, and Tuomas Aura. *Client-Side Vulnerabilities in Commercial VPNs*, pages 103–119. 11 2019. <https://arxiv.org/pdf/1912.04669.pdf>.
- [8] Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. Scaling instruction-finetuned language models, 2022.
- [9] Sarah Coble. Vpn usage in us quadruples. <https://www.infosecurity-magazine.com/news/vpn-usage-in-us-quadruples/>, 2020.
- [10] Natalie Cook, Susan Ayers, and Antje Horsch. Maternal post traumatic stress disorder during the perinatal period and child outcomes: A systematic review. *J Affect Disord.*, 2018(225):18–31, 2018.
- [11] Danilo Croce, Giuseppe Castellucci, and Roberto Basili. GAN-BERT: Generative adversarial learning for robust text classification with a bunch of labeled examples. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2114–2119, Online, July 2020. Association for Computational Linguistics.
- [12] Cyberwaters. Do streamers use a vpn (do you really need it?). 2021. <https://cyberwaters.com/do-streamers-use-vpn/>.
- [13] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805, 2018.
- [14] Agnieszka Dutkowska-Zuk, Austin Hounsel, Andre Xiong, Marshini Chetty, Nick Feamster, Molly Roberts, and Brandon Stewart. Practicing safe browsing: Understanding how and why university students use virtual private networks. *CoRR*, 2020.
- [15] Max Eddy. Why you need a vpn, and how to choose the right one. <https://in.pcmag.com/privacy/113633/what-is-a-vpn-and-why-you-need-one>, 2022.
- [16] Gunther Eysenbach and James E. Till. Ethical issues in qualitative research on internet communities. *BMJ*, 323(7321):1103–1105, November 2001.
- [17] Yahoo Finance. Global virtual private network (vpn) market to reach us\$77.1 billion by the year 2026. <https://finance.yahoo.com/news/global-virtual-private-network-vpn-122300132.html>, 2021.
- [18] Erza Gashi and Zhilbert Tafa. Permission-based privacy analysis for android applications. 10 2017. <https://knowledgecenter.ubt-uni.net/cgi/viewcontent.cgi?article=1230&context=conference>.

- [19] Margherita Grandini, Enrico Bagli, and Giorgio Visani. Metrics for multi-class classification: an overview, 2020.
- [20] Sumit Kumar Gupta and Dr. Dharendra Pratap Singh. Cbla: A clique based louvain algorithm for detecting overlapping community. *Procedia Computer Science*, 218:2201–2209, 2023. International Conference on Machine Learning and Data Engineering.
- [21] Emitza Guzman and Walid Maalej. How do users like this feature? a fine grained sentiment analysis of app reviews. *2014 IEEE 22nd International Requirements Engineering Conference, RE 2014 - Proceedings*, pages 153–162, 09 2014. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6912257>.
- [22] Gunnar Harboe and Elaine M. Huang. Real-World Affinity Diagramming Practices: Bridging the Paper-Digital Gap. In *Proc. CHI*, 2015.
- [23] Pengcheng He, Xiaodong Liu, Jianfeng Gao, and Weizhu Chen. Deberta: Decoding-enhanced bert with disentangled attention, 2021.
- [24] Nguyen Phong Hoang, Arian Akhavan Niaki, Jakub Dalek, Jeffrey Knockel, Pellaeon Lin, Bill Marczak, Masashi Crete-Nishihata, Phillipa Gill, and Michalis Polychronakis. How great is the great firewall? measuring china’s dns censorship. In *USENIX Security Symposium*, pages 3381–3398, 2021. <https://www.usenix.org/system/files/sec21-hoang.pdf>.
- [25] Hotstar. Watch tv shows, movies, specials, live cricket & football. <https://www.hotstar.com/>, 2023.
- [26] *Optuna: A hyperparameter optimization framework*.
- [27] Hulu. Hulu: Stream tv and movies live and online. <https://www.hulu.com/>, 2023.
- [28] Muhammad Ikram, Narseo Vallina-Rodriguez, Suranga Seneviratne, Mohamed Ali Kaafar, and Vern Paxson. An analysis of the privacy and security risks of android vpn permission-enabled apps. pages 349–364, 11 2016. <https://dl.acm.org/doi/10.1145/2987443.2987471>.
- [29] Asela Jayatilleke and Parakum Pathirana. Smartphone vpn app usage and user awareness among facebook users. pages 1–6, 10 2018. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8550081>.
- [30] Hamed Jelodar, Yongli Wang, Chi Yuan, and Xia Feng. Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78:15169–15211, 2017.
- [31] Chamandeep Kaur and Dr. Yogesh Sharma. The vital role of vpn in making secure connection over internet world. 8:2336–2339, 03 2020. <https://www.ijrte.org/wp-content/uploads/papers/v8i6/F8335038620.pdf>.
- [32] Mohammad Khan, Joe DeBlasio, Geoffrey Voelker, Alex Snoeren, Chris Kanich, and Narseo Vallina-Rodriguez. An empirical analysis of the commercial vpn ecosystem. pages 443–456, 10 2018. <https://dl.acm.org/doi/10.1145/3278532.3278570>.
- [33] Sheharbano Khattak, Mobin Javed, Syed Ali Khayam, Zartash Uzmi, and Vern Paxson. A look at the consequences of internet censorship through an isp lens. *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, pages 271–283, 11 2014. <https://dl.acm.org/doi/10.1145/2663716.2663750>.
- [34] Kristina. Why do i need vpn for gaming? 10 reasons it makes sense. 2023. <https://www.vpnmentor.com/blog/why-need-a-vpn-for-gaming/>.
- [35] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension, 2019.
- [36] Taoying Li, Jie Bai, Xue Yang, Qianyu Liu, and Yan Chen. Co-occurrence network of high-frequency words in the bioinformatics literature: Structural characteristics and evolution. *Applied Sciences*, 8:1994, 10 2018. <https://www.mdpi.com/2076-3417/8/10/1994>.
- [37] Eric Lim. app-store-scraper 0.3.5. <https://pypi.org/project/app-store-scraper/>, 2020.
- [38] Sui Luo, Heng Li, and Shuyuan Jin. Measuring information leakage of dns server. In *2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS)*, pages 796–805, 2021. <https://ieeexplore.ieee.org/abstract/document/9449155>.
- [39] Arunesh Mathur, Nathan Malkin, Marian Harbach, Eyal Péer, and Serge Egelman. Quantifying users’ beliefs about software updates. *CoRR*, abs/1805.04594, 2018. <http://arxiv.org/abs/1805.04594>.
- [40] Allison McDonald, Matthew Bernhard, Luke Valenta, Benjamin VanderSloot, Will Scott, Nick Sullivan, J. Alex Halderman, and Roya Ensafi. 403 forbidden: A global view of cdn geoblocking. In *Proceedings of the Internet Measurement Conference 2018, IMC ’18*, page 218–230, 2018.



- [41] P. V. Mockapetris. *RFC1034: Domain Names - Concepts and Facilities*. RFC Editor, 1987. <https://www.ietf.org/rfc/rfc1034.txt>.
- [42] Rebecca Moody. Vpn market report 2023: Who's got the biggest vpn market share? 2020. <https://www.comparitech.com/blog/vpn-privacy/vpn-market-share-report/>.
- [43] Moses Namara, Daricia Wilkinson, Kelly Caine, and Bart Knijnenburg. Emotional and practical considerations towards the adoption and abandonment of vpns as a privacy-enhancing technology. *Proceedings on Privacy Enhancing Technologies*, 2020:83–102, 01 2020.
- [44] Netflix. Netflix-watch tv shows online, watch movies online. <https://www.netflix.com/>, 2023.
- [45] Helmi Noman. Internet censorship and the intraregional geopolitical conflicts in the middle east and north africa. 01 2019. [bit.ly/36fvwG1](http://bit.ly/36fvwG1).
- [46] Vasile Perta, Marco Barbera, Gareth Tyson, Hamed Haddadi, and Alessandro Mei. A glance through the vpn looking glass: Ipv6 leakage and dns hijacking in commercial vpn clients. volume 1, 04 2015. <https://haddadi.github.io/papers/PETS2015VPN.pdf>.
- [47] PlanB. google-play-scraper 1.0.3. <https://pypi.org/project/google-play-scraper/>, 2020.
- [48] Mizanur Rahman, Nestor Hernandez, Ruben Recabarren, Syed Ishtiaque Ahmed, and Bogdan Carbunar. The art and craft of fraudulent app promotion in google play. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS '19*, page 2437–2454, New York, NY, USA, 2019. Association for Computing Machinery.
- [49] Reethika Ramesh, Leonid Evdokimov, Diwen Xue, and Roya Ensafi. Vpnalyzer: Systematic investigation of the vpn ecosystem. *Proceedings 2022 Network and Distributed System Security Symposium*, 2022.
- [50] Reethika Ramesh, Anjali Vyas, and Roya Ensafi. "all of them claim to be the best": Multi-perspective study of vpn users and vpn providers. *ArXiv*, abs/2208.03505, 2022.
- [51] Alan Ritter, Sam Clark, Mausam, and Oren Etzioni. Named entity recognition in tweets: An experimental study. In *Conference on Empirical Methods in Natural Language Processing*, 2011.
- [52] Punyajoy Saha, Binny Mathew, Kiran Garimella, and Animesh Mukherjee. "short is the road that leads from fear to hate": Fear speech in indian whatsapp groups. In *Proceedings of the Web Conference 2021, WWW '21*, page 1110–1121, 2021.
- [53] J. M. Saldana. *The coding manual for qualitative researchers*. SAGE, 2015.
- [54] Nissy Sombatruang, Tan Omiya, Daisuke Miyamoto, M. Angela Sasse, Youki Kadobayashi, and Michelle Baddeley. Attributes affecting user decision to adopt a virtual private network (VPN) app. *CoRR*, abs/2008.06813, 2020. <https://arxiv.org/abs/2008.06813>.
- [55] Fateme Tahmasbi, Leonard Schild, Chen Ling, Jeremy Blackburn, Gianluca Stringhini, Yang Zhang, and Savvas Zannettou. "go eat a bat, chang!": On the emergence of sinophobic behavior on web communities in the face of covid-19. *Proceedings of WWW*, 2021.
- [56] Dallas Thomas. 20 privacy & security settings you need to check on your google pixel. 2019. [bit.ly/36nAtN9](http://bit.ly/36nAtN9).
- [57] Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. Llama 2: Open foundation and fine-tuned chat models, 2023.
- [58] Jack Wilson, David McLuskie, and Ethan Bayne. Investigation into the security and privacy of ios vpn applications. pages 1–9, 08 2020. <https://dl.acm.org/doi/10.1145/3407023.3407029>.
- [59] Jan Youngren. Who is dominating the rising vpn market in 2023? 2023. <https://vpnpro.com/blog/vpn-market-share-overview/>.
- [60] Qing Zhang, David Y. Wang, and Geoffrey M. Voelker. Dspin: Detecting automatically spun content on the web. In *Network and Distributed System Security Symposium*, 2014.