SecurePost : Verified Group-Anonymity on Social Media

Michael Nekrasov
UC Santa Barbara
mnekrasov@cs.ucsb.edu

Daniel Iland
UC Santa Barbara
iland@cs.ucsb.edu

Miriam Metzger
UC Santa Barbara
metzger@comm.ucsb.edu

Ben Zhao
UC Santa Barbara
ravenben@cs.ucsb.edu

Elizabeth Belding
UC Santa Barbara
ebelding@cs.ucsb.edu

Abstract

As Internet freedoms are increasingly threatened both at home and abroad, marginalized groups, such as journalists, activists, and government watchdogs require new tools to retain free and open discourse on-line. In this paper, we introduce SecurePost - a tool for verified group anonymity on social media. SecurePost gives social media posters anonymity while safeguarding group credibility through the use of revocable asymmetric keys and an anonymizing proxy. It provides trust to readers via the use of HMAC verification signatures appended to posts verifying integrity and authenticity of a post. We root our work in survey-based research and ethnographic interviews conducted with marginalized groups in Mongolia, Turkey, and Zambia from 2014 to 2016. SecurePost widens the toolkit of security applications, by giving vulnerable communities a way of balancing individual anonymity and safety with group credibility.

1 Introduction

Worldwide, the Internet is a critical medium for the exchange of ideas. Due to its importance, the United Nations has declared Internet access and free speech basic human rights [34, 35]. However, Internet freedoms are under attack globally [23, 32, 7, 18, 15, 19]. Even liberal democracies increasingly restrict content [3, 13]. By limiting the ability for groups to organize and share ideas, governments and corporations are able to suppress voices that are critical to a functioning democracy. Limiting Internet freedom limits personal freedom.

Censorship of mainstream social media is a challenge not adequately addressed by existing counter-censorship tools. On the one hand, groups seeking to broadcast their ideas to the widest possible audience prefer social media as the platform for communication due to its pervasiveness. On the other hand, out of fear of reprisal, individuals in a group may wish to retain anonymity while maintaining credibility. These two requirements provide a need that existing tools do not adequately address. We cater to these requirements through SecurePost, a novel tool for verified group-anonymity on social media. SecurePost developed through research and partnerships with affected individuals as a means of balancing personal anonymity with group credibility. In this paper we present the research that went into developing SecurePost, its technical contributions and operation, and an evaluation of our work.

2 Research into Internet Freedom

As the basis for developing SecurePost, we partnered with communities directly affected by censorship. We sought to design a tool that would augment the communication of these communities while paying heed to restrictions imposed by needs and existing usage patterns. To that end, we conducted ethnographic interviews and survey-based research from 2014 to 2016 in three distinct communities: Ulaanbaatar, Mongolia; Istanbul, Turkey; and Lusaka, Zambia.

We selected these countries as a cross-section into global censorship. They are in differing levels of socio-economic development, with disparate cultural and historical backgrounds, leading to a variety of censorship strategies. Moreover, the research examining censorship for these countries is sparse compared to China [39, 17, 6, 37] and the Middle East [2, 4, 21]. We believe our work could bring new perspectives to the discussion of Internet censorship.

In total, across the three countries we surveyed 525 individuals, conducted 109 formal interviews, as well as informal conversations with dozens more. We obtained IRB approval prior to conducting the user studies. We predominantly interacted with higher socio-economic status participants compared to the overall populations. In the interviews we focused on individuals vulnerable to Internet censorship. They included journalists, politi-

cal groups, LGBT activists, government watchdogs, academics, and other individuals vulnerable to persecution and censorship. All three countries have harsh laws limiting on-line content with frequent examples of censorship [31, 10, 11, 16, 28, 26].

We sought to understand how individuals and groups are affected by Internet censorship, and the limitations of currently available security tools. Our team included experts from computer science, communication, and film and media studies. We used descriptive statistics of survey data coupled with ethnographic analysis of interviews to guide initial design. We coupled this with iterative critique from participants during the development of SecurePost to guide design decisions. For brevity, this paper focuses on research outcomes related to verified group anonymity. We provide a more comprehensive exploration of insights from our research on Internet freedom in [22].

Our research focused on nation capitals, which have a far greater adoption of technology than the rural countryside. Capitals are also epicenters of journalism and political activism, as well as censorship battlegrounds. Our results confirm global trends in these local contexts. The majority of participants in our research used smartphones, usually running Android, as the primary means of accessing the Internet. This is in line with global statistics, where, as of 2016, 49% of the world's population had a mobile-broadband subscription, while only 12% had fixed broadband subscriptions [24]. In terms of operating system, as of end of 2016, Android made up 82% of the market share [12]. We therefore targeted mobile devices running Android for development.

The majority of our participants used social media to communicate on-line, with Facebook the most used online social network in all three countries. Globally, as of June 2017, Facebook had approximately 2 billion active users [27], the most of any social media platform. The majority of our participants reported limiting their social media usage and self-censoring content due to feeling unsafe on-line. For participants, the use of purpose-built social media platforms for anonymity, such as [5], would be unsuitable as the expansive active user base of top social media sites provides a substantially broader audience than other platforms. We therefore selected Facebook and, due to its similarity and high usage rate, Twitter, as the primary social media platforms to support.

A key frustration voiced by our participants is the clash between anonymity and trust. The public relies on the reputation of an account or organization to produce trustworthy content. However, reputable accounts with self-identifying information are targets for oppression [9, 20]. To avoid prosecution (or worse), users sometimes elect to post anonymously. But, anonymity comes at a cost. For example, some journalists we interviewed at the Zambian Watchdog (ZWD) indicated that they maintain anonymous blogs to avoid being monitored by government authorities when discussing sensitive issues. Yet others object to the use of anonymity and lack of a physical address for accountability and verification, which they say has led the ZWD to lose credibility among some readers.

Anonymous accounts can get lost in the noise, finding it difficult to establish trust. Increasingly, adversaries generate spam [36] and fake-news [1] to drown out competing ideas. Journalists whose livelihood depends on reputation find it difficult to resolve anonymous communication and reputation building.

Additionally, masking public identity of an account is not enough, as social media platforms are inundated with information requests and may co-operate with governments or corporations to reveal IPs of users and tie identities to individual posts [33, 8]. Requiring all users to use VPNs or Tor [29] for each post is often unrealistic as groups can be composed of posters with varying technological expertise, and a single mistake can be costly. In more extreme cases, adversaries confiscate devices allowing access to their applications [30].

3 Verified Group-Anonymity

Based on participant feedback, we designed SecurePost to protect groups against some of the unmet censorship challenges outlined in Section 2. While individuals can communicate privately and anonymously amongst themselves using a variety of existing security tools, groups encounter difficulties communicating publicly when confronted with censorship. SecurePost seeks to address this by providing a mechanism for verified group-anonymity for mainstream social media. We provide its technical description in Section 4.

SecurePost allows individual members of groups to share a single group identity, while retaining individual anonymity. When a member of a SecurePost group posts on social media, the identity of the poster is not tied to the post. The identity is hidden from the network, which may be monitored by government, corporations, or other adversaries. The identity is also hidden from other members of the group. So, the group itself has a social media presence that builds trust, while adversaries are unable to link specific individuals to a particular post, giving individuals plausible deniability.

Because preserving credibility is important, SecurePost allows compatibility with existing social media accounts. So, users may choose to form groups based on existing organizations. For example, SecurePost allows a newspaper, like The Zambian Post, to use their name and established reputation on Twitter while protecting individual reporters.

When a higher level of anonymity is required, users may choose to form new groups, not linked to a physical location or known affiliation. These groups can build credibility together while retaining control of group-management. SecurePost allows an invitation scheme where no individual group member knows the complete membership roster or even the number of members in the group. Groups starting afresh are able to use SecurePost to hide membership and affiliation. This gives greater protection when the group is targeted by adversaries.

Social network accounts can be seized, hacked, or infiltrated. Compromised accounts can be used to post disinformation or to edit existing content. SecurePost offers a tool for signing and verifying posts via the use of cryptographic signatures. If users choose to enable this feature, every post is signed with a HMAC verifying that the post came from an approved poster who was invited into the group. The HMAC also ensures the post was not modified in any way. Even if the social network account was hacked, without being invited into the group through the app, the attacker is unable to forge new messages or edit existing ones. Attackers are limited to deleting posts or locking out the account.

Through verified group anonymity, users are able to form groups that build trust. By inviting sources they trust, groups can build a web of trust without revealing the identities of their members to themselves or others. Members of the public can easily see what posts are certified by the group rather than tampered with or forged. If the trust of a group is compromised, SecurePost allows trusted administrators to boot all members from the group, invalidate previous posts, and restart the group creation process anew.

4 SecurePost

SecurePost comprises three coordinated modules, which provide group-anonymity and verified authenticity. Group members post via an Android application. These posts are relayed through a proxy server, to social media where they appear publicly. Optionally, users may verify the authenticity of these posts through a browser plug-in, which automatically checks applicable posts for authenticity.

Currently SecurePost supports only Twitter and Facebook. However, the code is designed to be modular, so any platform providing an API for posting could be incorporated in the future. We discuss the details of these modules in this section.

4.1 The SecurePost Android Application

The primary module of SecurePost is the mobile application. We support Android API level 10, which includes

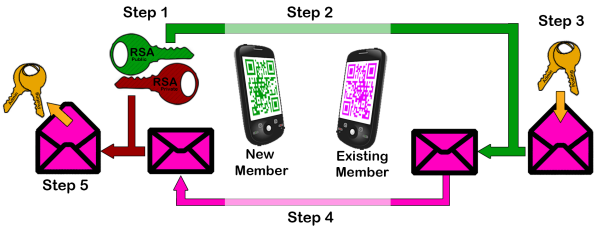

Figure 1: Process flow of visual invitation scheme. Step 1: generate asymmetric key pair. Step 2: public key sent as QR code. Step 3: public key encrypts group credentials. Step 4: send encrypted invite as QR code. Step 5: private key decrypts group credentials.

devices running Android 2.3.3 and above. As of May 2017, this accounts for 99.9% of Android devices registered with Google [14]. Members of a group use the Android application for posting anonymously to social media. Through the app, users can form new groups, manage group membership, and view posts to their groups. Viewing the posts does not require using the application; anyone can view the posts on social media.

4.1.1 Creating a New SecurePost Group

To link a social media account to SecurePost, users use the platform specific login API displayed through the app. This generates an access token, which is sent to the proxy server and discarded by the app. This is the only time the app requires any user to enter the social media login credentials. To protect the group creator, we recommend completing this step in conjunction with an anonymity service, such as Tor [29] or a VPN, running on the device. Once the group is created, the current user becomes the group administrator. Users can set up an unlimited number of groups.

4.1.2 Inviting Others

Typically, organizations share a social media account by sharing a password. This leads to problems tracking authorization and integrity of the account. SecurePost instead relies on public key cryptography.

When creating an account the app generates two asymmetric 2048 bit RSA key pairs. The app stores the private keys, and transmits the public keys to the proxy server. One pair of keys denotes the app's posting credentials, and the other pair denotes the administrative credentials.

Group members can invite others. Inviting a new member requires that the existing member's app passes the credentials to the recruit in the form of RSA private keys. Inviting new administrators clones both administrative and posting keys. Inviting contributors, with no administrative rights, clones only the posting key.

The complexity of the credentials exchange is hidden from the users in the form of a short and simple invite wizard. We provide two methods to join: an optical face-to-face QR-Code exchange and a remote invitation.

Face-to-Face Credential Exchange: The more secure invitation method is a two step QR code exchange, depicted in Figure 1. The recruit shows an ephemeral 2048 bit RSA public key in the form of a scanable QR code. The existing member scans this key and uses it to encrypt the group credentials (the group's private keys) and display the encrypted invite back in the form of a new QR code. The recruit scans this QR code. The app decodes the group credentials using the ephemeral private key. The recruit is now part of the group, able to authenticate with the proxy, and ready to post. By using a two step process, an adversary visually observing the exchange would be unable to decrypt the group credentials without the recruit's ephemeral private key.

Alternative Credential Exchange: As face-to-face invitation is not possible in all contexts, we allow users to share the private keys by encoding them into an invite code, shareable via the Android share intent out of band. Recruits can join by copying the invite code into their SecurePost app. To prevent adversaries from infiltrating the group, we recommend sharing through secure end-to-end encrypted applications.

Initial Design: In our initial designs, group authentication was achieved through a time synchronized hash chain for message authentication using a long passwords as a seed. We envisioned short-lived groups formed at social action events, such as protests. During user testing, participants explained the importance of long lasting groups that build trust and credibility over time. They also struggled with remembering and sharing complex passwords. This led us to the multi-use public key solution as a means of credential exchange, which is an intuitive and more secure method of invitation. The same credentials used for server authentication can be re-used for post verification as described in section 4.1.4.

4.1.3 Group Management

The invitation scheme intentionally lacks a user registry. To the proxy server, it appears as if a single user is accessing the system. Membership of a group is only known out of band, through the social interactions between people. No one necessarily knows all members of a group or even how many members there are. Group members can enlist confidential contributers without revealing identities to the rest of the group.

This approach is advantageous for an individual's secrecy and plausible deniability. It also precludes the ability to ban a specific member from the group. However, as groups could become compromised or trust in a group member may be misplaced, we provide a mechanism for group control. Any administrator (i.e. anyone with possession of the administrative private key) can reset the group. This re-issues a new pair of keys to which only the user performing the reset has access. All other members and previous posts become invalidated, and must be re-invited with greater care.

4.1.4 Posting

Any user with a valid private key can post using SecurePost. Unlike other social media clients, all posts in a given SecurePost group appear to come from one social media account, which was linked at group creation. No one inside or outside the group can identify which group member composed a given post. This provides the foundation for group anonymity of our application.

To prove membership, posts are signed with the group's posting private key. The key is used to compute an HMAC, which serves as a signature providing both a proof of integrity and authenticity. For group members, this signature is automatically checked by the application when displaying messages. Those reading the message directly on the social media platform can use our Browser Extension, described in Section 4.3, to authenticate the message.

When the group is reset, all previous messages using the old key are marked invalid. Users can still see these messages, but they are visually flagged and come with an explanatory disclaimer.

Posting Multimedia: Due to demand from our users, we also added the ability to post images. We are contemplating adding video and audio posts as well. While these posts benefit from group anonymity, they do not support signing for validity and authenticity. This is because social-media platforms compress and alter images and other media uploaded to their platforms, which invalidates the HMAC. We mark multimedia messages as unverified when displaying them to users.

4.1.5 Storage

All group memberships, keys, and posts are stored in a SQLCipher [38] encrypted database. When the app is first installed, users choose a long master password for unlocking the application. Each time the application is started, this password is needed to decrypt the database. While the application is running it spawns a notification and can be sent to the background while preserving the unlocked state. Dismissing the notification or exiting the application in another way re-encrypts the database. Losing this password is unrecoverable. Users who lose the password are prompted to reset the application and must ask to rejoin all their groups.

Our app is tailored to populations that run older operating systems on cheap devices, as discussed in Section 2, which do not support newer security features like full disk encryption or a secure enclave. The encrypted database provides an added layer of security for these devices. Even if the phone is confiscated by an adversary, the adversary would need to perform a costly attack to decrypt data, which would still not expose group membership.

Self-Destruct Password: In addition to the application password, we allow an optional "self-destruct" password. In the event the device is confiscated and the owner is under duress, they can enter or give out this false password. Entering it in the application irretrievably wipes the content.

4.2 The SecurePost Proxy Server

The app works in conjunction with a proxy server of our design. Social media platforms can log IPs of devices and identify users even if they are using the same access tokens. These platforms can then be pressured by governments to give up access logs and help identify users. The proxy server is therefore needed as an intermediary between the SecurePost application and social media. In addition, the proxy server, and not the application, stores social media access tokens. If a group is reset, the token remains consistent on the proxy server. Only the public keys that are used for user verification are updated.

In case the server is compromised, we take steps to limit its power. The server does not log IPs or keep any metrics about users. The server also does not have access to the private keys used to sign posts. If an adversary gained access to the server, they could compose posts but not sign them. The server also does not have access to the login information for a social media platform. If the server was compromised and began posting erroneous posts, the posts would still show as invalid. The owner of a social media account password could login to the platform and revoke the access token.

The server consists of a standalone Java application running Jetty coupled with a MongoDB No-SQL Database. A JSON REST API running over HTTPS interfaces the app and server. For scalability, the app and database can run on independent servers.

We run a cloud-based instance of the server code for users of our applications on Amazon Web Services. However, we do not expect users to trust the server we provide. Therefore, the code for the server and application is open source and we make it straight-forward to configure the application to point to a different proxy server instance.

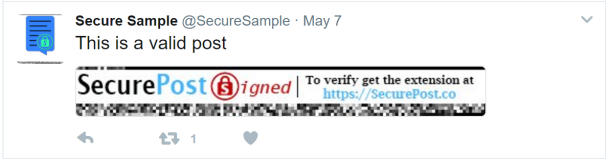

Figure 2: Sample signature used by SecurePost.

4.3 The Browser Extension

For all social media users, including those that do not have posting rights to the group, we provide a browser extension that automatically verifies a post's signature. While we officially support it only for Chrome, it is also compatible with Firefox and Opera.

4.3.1 Independent Post Signatures

Because we envision users may choose to use their own proxy server, we wanted to make the extension independent of the proxy server. This way, the browser extension can verify messages of any group, using any custom proxy server, provided they use the same verification scheme. As the number of users verifying posts would be magnitudes higher than post creators, this also reduces server load and hardware requirements.

If the SecurePost group enables the option to use verification, each post on the social media platform is appended with an image like the one in Figure 2. Since social media platforms compress images, we use a compression resistant encoding for the HMAC signature. The HMAC is encoded as a series of monochromatic 9 pixel (3x3) squares. The browser extension reads the squares and decodes the signature back into binary. This technique does not require co-operation with the social media platform or the proxy server and bypasses the character limit on platforms like Twitter.

4.3.2 Providing the Public Key

To verify the HMAC, the extension requires the posting public key. In the same scheme as the signature, when the group is created, the proxy encodes the public key into the profile image on Twitter and the cover image on Facebook. The signature appears as a thin strip of monochromatic blocks at the bottom of the profile image.

4.3.3 Verifying Posts

When a social media feed or account is displayed, the browser automatically verifies posts. If the account uses SecurePost, as identified by a pixel pattern encoded in the corner of the profile image, the extension validates the posts. It marks them appropriately - using a mix of color and symbols as shown in Figure 3. Posts lacking

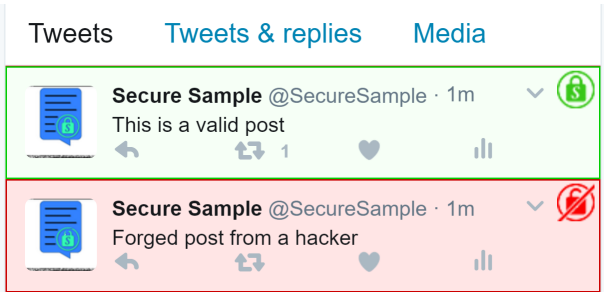

Figure 3: Example of post verification on Twitter.

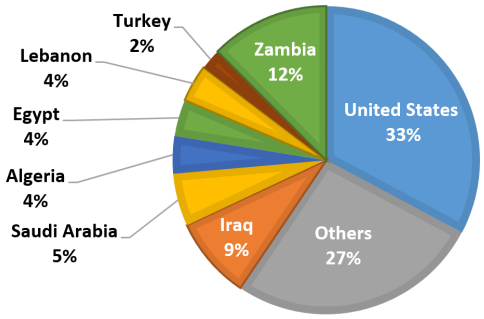

Figure 4: SecurePost application installation by country.

signatures (i.e. those posted without using the app) and non text-based posts are not verifiable and are marked accordingly. The plug-in also automatically hides signature images in post bodies from the user, reducing the visual overhead of embedding images in every verified post.

Initial Design: In our initial design, we experimented with text-based signatures using 21-bit CJK Unified Ideographs. These characters allow for widely supported high bit-density character encoding, while minimizing the impact of social media character limits. Unfortunately, in user testing in Mongolia, participants expressed that this approach was unsuitable due to the social impact of appearing to affiliate with China or Korea. Users had no such issues with image-based encoding.

5 Evaluation

Unlike most security tools, we built SecurePost from the ground up, in partnership with affected communities. Throughout the design and refinement process, we received iterative feedback from our partners. We limited our design space to solutions that work with existing uncooperative systems utilized by individuals with only a modest level of technical expertise. These restrictions have, however, led to novel technical solutions comprising a tool that is accessible to the average user.

We have published our app free of charge to the Google Play store, and the verifier to the Google Chrome store. Our app is available in 7 languages (Arabic, English, French, Mongolian, Russian, Spanish, Turkish). We also provide our source code for the app, server, and browser extension on our website [25]. We are currently working with our partners to test and evaluate the finished version of our tool.

Due to the anonymity constraint, we do not collect usage data beyond what is necessary for functionality of the app and server. Based on data collected from the Google Play store, as of June 2017 we have had 375 installs of our app. Users of our app reside in the countries we were explicitly targeting: USA, Mongolia, Turkey, Zambia, as well as other countries where censorship is a problem as shown in Figure 4. Users of our app have created 66 groups and made 335 posts. From remote correspondence and in-person training with our partners, we have received positive feedback on our work. Additionally, our app prompts for anonymous feedback after a week of usage; we are still collecting that data.

6 Discussion

SecurePost addresses a set of unmet needs for groups. By providing group anonymity, users are able to build reputation as a collective without exposing individuals. Even if adversaries collude with a social media platform, IPs and identities of group members are not leaked. Social media accounts can be shared without giving up the account passwords and without relinquishing total control. If the social media account is seized or hacked, without the cryptographic signature, the verification system can identify fraudulent posts. Using SecurePost, users can strike a balance between anonymity and reputation.

SecurePost users can post with greater confidence. Unfortunately, like other anonymity applications, users must still be mindful of the content of their posts. Revealing personal information in the content of a message can identify the individual poster. Additionally, adversaries with a sufficient view of the network may still implement de-anonymization through timing analysis. We hope to address this vulnerability in future work.

A common concern of security applications is the potential misuse by ill-meaning organizations, like terrorist cells. Because all the data is posted publicly to social media, our app does not expand the capabilities of malevolent secret communication. While our tool allows users to remain anonymous, it does not prevent social media platforms from blocking the account as a whole. We push the burden of deciding whether a group is dangerous to the social media platform.

Our work provides insight into working with communities to incorporate security into commonly used services. By encoding cryptographic information alongside text, we present a novel method of adding security to non-cooperative social media platforms.

References

[1] ALLCOTT, H., AND GENTZKOW, M. Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives 31*, 2 (2017), 211–236.

[2] ARYAN, S., ARYAN, H., AND HALDERMAN, J. A. Internet censorship in Iran: A first look. In *Proceedings of the 3rd USENIX Workshop on Free and Open Communications on the Internet* (August 2013).

[3] BREINDL, Y., AND WRIGHT, J. Internet filtering in liberal democracies. In *Proceedings of the 2nd USENIX Workshop on Free and Open Communications on the Internet* (Bellevue, WA, August 2012), USENIX.

[4] CHAABANE, A., CHEN, T., CUNCHE, M., DE CRISTOFARO, E., FRIEDMAN, A., AND KAAFAR, M. A. Censorship in the wild: Analyzing internet filtering in Syria. In *Proceedings of the 2014 Conference on Internet Measurement Conference* (2014), IMC '14, pp. 285–298.

[5] CORRIGAN-GIBBS, H., BONEH, D., AND MAZIERES, D. Riposte: An anonymous messaging system handling millions of users. In *2015 IEEE Symposium on Security and Privacy* (May 2015), pp. 321–338.

[6] CRANDALL, J. R., ZINN, D., BYRD, M., BARR, E. T., AND EAST, R. Conceptdoppler: a weather tracker for internet censorship. In *ACM Conference on Computer and Communications Security* (2007), pp. 352–365.

[7] DAINOTTI, A., SQUARCELLA, C., ABEN, E., CLAFFY, K. C., CHIESA, M., RUSSO, M., AND PESCAPÉ, A. Analysis of country-wide internet outages caused by censorship. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference* (2011), IMC '11, pp. 1–18.

[8] FACEBOOK. Government requests report. `https://govtrequests.facebook.com/`. (Accessed May. 2017).

[9] FIDH. Turkey: Provisional release of human rights lawyer Mr. Levent Piskin. `https://www.fidh.org/en/issues/human-rights-defenders/turkey-provisional-release-of-human-rights-lawyer-mr-levent-piskin`, Nov 2016.

[10] FLORIO, A. D., VERDE, N. V., VILLANI, A., VITALI, D., AND MANCINI, L. V. Bypassing censorship: A proven tool against the recent internet censorship in Turkey. In *2014 IEEE International Symposium on Software Reliability Engineering Workshops* (Nov 2014), pp. 389–394.

[11] FREEDOM HOUSE. Mongolia: Freedom of the press 2016. `https://freedomhouse.org/report/freedom-press/2016/mongolia`, 2016.

[12] GARTNER. Gartner says worldwide sales of smartphones grew 7 percent in the fourth quarter of 2016. `https://www.gartner.com/newsroom/id/3609817`, Feb 2017.

[13] GOLDMAN, D. Donald Trump wants to 'close up' the internet. `http://money.cnn.com/2015/12/08/technology/donald-trump-internet/`, Dec 2015.

[14] GOOGLE. Android developers dashboard. `https://developer.android.com/about/dashboards/index.html`. (Accessed on 05/14/2017).

[15] KELLY, S., EARP, M., REED, L., SHAHBAZ, A., AND TRUONG, M. Privatizing censorship, eroding privacy. `https://freedomhouse.org/sites/default/files/FH_FOTN_2015Report.pdf`, Oct 2015.

[16] KELU, K. The plight of the Zambian watchdog: Embattled opposition news site goes down. `https://advox.globalvoices.org/2016/10/11/the-plight-of-the-zambian-watchdog-embattled-opposition-news-site-goes-down/`, October 2016.

[17] KING, G., PAN, J., AND ROBERTS, M. E. How censorship in China allows government criticism but silences collective expression. *American Political Science Review 107*, 2 (2013), 326–343.

[18] LEE, T. B. Here's how Iran censors the Internet. `https://www.washingtonpost.com/news/the-switch/wp/2013/08/15/heres-how-iran-censors-the-internet`, Aug 2013.

[19] LIM, K., AND DANUBRATA, E. Singapore seen getting tough on dissent as cartoonist charged. `http://www.reuters.com/article/us-singapore-dissent-idUSBRE96P0AF20130726`, Jul 2013.

[20] LOWEN, M. Is Gollum good or evil? Jail term in Turkey hinges on answer. `http://www.bbc.com/news/world-europe-32302697`, Apr 2015.

[21] NABI, Z. The anatomy of web censorship in Pakistan. In *Proceedings of the 3rd USENIX Workshop on Free and Open Communications on the Internet* (August 2013).

[22] NEKRASOV, M., PARKS, L., AND BELDING, E. Limits to internet freedoms: Being heard in an increasingly authoritarian world. In *Proceedings of the Third Workshop on Computing Within Limits* (June 2017), ACM LIMITS '17.

[23] PETERSON, A. Turkey strengthens Twitter ban, institutes IP level block. `https://www.washingtonpost.com/news/the-switch/wp/2014/03/22/turkey-strengthens-twitter-ban-institutes-ip-level-block`, March 2014.

[24] SANOU, B. ICT facts and figures 2016. `http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2016.pdf`, 2016.

[25] SECUREPOST. Securepost - safe, secure, social media. `https://securepost.co`. (Accessed on May 2017).

[26] SHAHEEN, K. Turkey arrests pro-Kurdish party leaders amid claims of internet shutdown. `https://www.theguardian.com/world/2016/nov/04/turkey-arrests-pro-kurdish-party-leaders-mps`, November 2016.

[27] STATISTA. Global social media ranking 2017. `https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/`, April 2017.

[28] STOCKHOLM CENTER FOR FREEDOM. Jailed and wanted journalists in Turkey. `http://stockholmcf.org/updated-list/`. (Accessed on 05/14/2017).

[29] SYVERSON, P., DINGLEDINE, R., AND MATHEWSON, N. Tor: the second generation onion router. In *Proceedings of the USENIX Conference on Security Symposium. USENIX Association* (Berkeley, CA, USA, 2004), SSYM'04, USENIX Association, pp. 21–21.

[30] TAYLOR, A. This single tweet got a Turkish journalist detained. `https://www.washingtonpost.com/news/worldviews/wp/2014/12/30/this-single-tweet-got-a-turkish-journalist-detained/?utm_term=.234e55c52105`, December 2014.

[31] TUMFWEKO. Wina justifies beating of Komboni radio owner. `https://www.tumfweko.com/2016/10/09/wina-justifies-beating-of-komboni-radio-owner/`, Oct 2016.

[32] TURKEY BLOCKS. New internet shutdown in Turkey's Southeast: 8% of country now offline amidst Diyarbakir unrest. `https://turkeyblocks.org/2016/10/27/new-internet-shutdown-turkey-southeast-offline-diyarbakir-unrest/`, Oct 2016.

[33] TWITTER. Information requests. `https://transparency.twitter.com/en/information-requests.html`. (Accessed on May. 2017).

[34] UNITED NATIONS. Universal declaration of human rights. http://www.un.org/en/universal-declaration-human-rights/, December 1948. (Accessed on 02/16/2017).

[35] UNITED NATIONS. The promotion, protection and enjoyment of human rights on the internet. https://www.article19.org/data/files/Internet_Statement_Adopted.pdf, June 2016.

[36] VERKAMP, J.-P., AND GUPTA, M. Five incidents, one theme: Twitter spam as a weapon to drown voices of protest. In *Proceedings of the 3rd USENIX Workshop on Free and Open Communications on the Internet* (August 2013).

[37] WINTER, P., AND LINDSKOG, S. How the great firewall of China is blocking Tor. In *Proceedings of the 2nd USENIX Workshop on Free and Open Communications on the Internet* (August 2012), USENIX.

[38] ZETETIC. Sqlcipher. https://www.zetetic.net/sqlcipher/sqlcipher-for-android/. (Accessed on 05/13/2017).

[39] ZITTRAIN, J., AND EDELMAN, B. Internet filtering in China. *IEEE Internet Computing 7*, 2 (Mar 2003), 70–77.