Using Internet Measurements to Map the 2022 Ukrainian Refugee Crisis

Tal Mizrahi
Technion — Israel Institute of Technology

Jose Yallouz
Technion — Israel Institute of Technology

ABSTRACT
The conflict in Ukraine, starting in February 2022, began the largest refugee crisis in decades, with millions of Ukrainian refugees crossing the border to neighboring countries and millions of others forced to move within the country. In this paper we present an insight into how Internet measurements can be used to analyze the refugee crisis. Based on preliminary data from the first two months of the war we analyze how measurement data indicates the trends in the flow of refugees from Ukraine to its neighboring countries, and onward to other countries. We believe that these insights can greatly contribute to the ongoing international effort to map the flow of refugees in order to aid and protect them.

1 INTRODUCTION
The 2022 war in Ukraine, which started with Russia’s invasion on 24 February 2022, caused Europe’s largest refugee crisis since World War II [15]. The unprecedented refugee crisis has been an increasing concern, with nearly 6 million refugees exiting Ukraine as of May 2022 [21] and over a quarter of the Ukrainian population being displaced [4] inside the country. The United Nations High Commissioner for Refugees (UNHCR) is chartered to aid and protect refugees in this crisis.

The UNHCR publishes data [21] about refugees that exited Ukraine to one of its seven neighboring countries on a daily basis. The accumulated number of refugees since the beginning of the war is illustrated in Figure 1a, and a histogram of the neighboring countries to which the refugees crossed the border is shown in Figure 1b.

One of the first steps in aiding these refugees is being able to monitor the distribution of refugees throughout the world. However, while the UNHCR continuously keeps track of the influx from Ukraine to the neighboring countries, once exiting Ukraine individuals may move freely between countries and therefore it is difficult to have a clear picture of the refugee distribution throughout the world, and especially throughout Europe since border crossing between EU countries is not monitored.

Preliminary findings about the impact of the war on the Internet performance in Ukraine have been discussed in [1, 2]. The dramatic differences between the Internet performance changes in Ukraine and those in Russia at the beginning of the conflict are presented in [14]. Clearly the connectivity during the war was affected by infrastructure aspects such as power outages or damaged communication lines and equipment, causing routing changes and congestion along bottlenecks, and in some cases resulting in user performance degradation. We argue that the refugee crisis was also a major factor in some of these performance changes.

In this paper we use Internet measurement data from multiple sources to show that the Ukrainian refugee crisis affected specific Internet performance metrics not only in Ukraine, but in other countries as well. We demonstrate that during the first 2-3 weeks of the conflict various performance metrics such as the average traffic rate and the mobile device usage significantly changed, having a clear correlation to the flow of refugees.

Our analysis shows how publicly available Internet measurement data can be used to analyze the geographic distribution and the flow of refugees over time. We use website analytics of Ukrainian sites in order to provide a maximum likelihood estimation of the Ukrainian presence in each country. Our analysis is based solely on publicly available information, and produces large-scale statistics that do not compromise privacy aspects. We believe that our approach can be used to aid the ongoing effort in mapping the refugee crisis in order to help and support the refugees.

The rest of this paper is organized as follows. Section 2 describes the sources of data used in this paper. The impact of the refugee crisis on the Internet performance in Ukraine and the neighboring countries is discussed in Section 3, and
Section 4 introduces a measurement-based mapping of the refugee crisis. Related work is discussed in Section 5, and concluding remarks are presented in Section 6.

2 DATA SOURCES

2.1 UNHCR
The United Nations High Commissioner for Refugees (UNHCR) publishes daily updates [21] about the number of refugees from Ukraine arriving to each of Ukraine’s neighboring countries. This data is based on information provided to the UNHCR by authorities from border crossing points. As noted above, while these statistics provide information about Ukrainians crossing the border to neighboring countries, individuals may move freely within European countries, and therefore it is estimated that a large number of people have moved onwards to other countries. The data in this paper was extracted from [21] on the second week of May 2022.

2.2 Statcounter
Statcounter [19] is a web analytics tool that tracks information about web page views collected from over 2 million sites globally. Statcounter allows flexible extraction of data on a per-country basis including, for example, web browser or social network statistics. The statistics we used in the current work include mobile and desktop usage rates, mobile vendor usage rates, and web search engine rates. Most of this data was analyzed over a period of 3 months, with some of the data analyzed over a period of several years.

2.3 Similarweb
Similarweb [17] publishes various website analytics data. In the current paper we used Similarweb’s ranking of the 50 top Ukrainian websites, as well as the number of visits per month to each website we analyzed. This data was accessed on 16 May 2022.

2.4 Google Transparency Report
Google’s transparency report [11] includes continuously updated data about traffic rates. The site provides data on a per-country basis for each of Google’s products such as Web Search and YouTube. The data provides the ratio between the country’s request rate and the worldwide request rate for each of Google’s products. In this paper we used data over a period of 3 months. For each day in the measurement period we extracted the peak normalized traffic rate, allowing to analyze the trends of the normalized traffic rate in each of the countries we analyzed.

2.5 Cloudflare Radar
Cloudflare Radar [7] is a site that presents detailed data about Cloudflare’s traffic. The site enables per-country filtering, including per-country normalized traffic rate over the last 60 days. We used this sliding window of 60 days to keep track of the traffic rates in Ukraine and six of its neighboring countries for a period of 3 months. For each day during this period we extracted the normalized peak traffic rate. Cloudflare also published website visit rate on a per-country basis, which was used in our analysis of Ukrainian websites. This data was accessed on 16 May 2022.

2.6 RIPE Atlas
RIPE Atlas [16] is a global measurement platform that uses over 11000 probes spread throughout the world. Measurements are continuously performed and published, as well as the status of each of the probes. We analyzed data about the connectivity status of about 200 probes in Ukraine over a period of a year.
2.7 Speedtest

Speedtest [18] by Ookla is one of the most commonly used sites for web-based Internet performance testing. Speedtest result statistics are published monthly [18] on a per-country basis, including Speedtest’s Global Index ranking; each country is ranked according to the median download speed. A separate ranking is published for fixed and for mobile tests. In this paper we used Speedtest data to analyze how the ranking of Ukraine and each of the neighboring countries changed since the beginning of the conflict.

3 IMPACT OF THE REFUGEE CRISIS ON INTERNET PERFORMANCE

3.1 Performance in Ukraine

As previously discussed, the Internet performance in Ukraine after the beginning of the conflict was affected by infrastructure damage and outages. However, as we show in this section, some of the performance metrics and behavior were significantly different in the first 2-3 weeks of the war due to the large flow of refugees, as well as vast population displacement within Ukraine.

The rate of refugees exiting Ukraine according to [21] is illustrated in Figure 2a, with the first 3 weeks marked by dotted lines, emphasizing the high rate during this period. Figures 2b-2d illustrate the unusual performance during this period of time, showing a high correlation with the refugee rate, which suggests that the most significant factor that affected these performance changes during this 3 week period is the refugee crisis. The figures throughout this paper are marked with a dotted line, indicating the beginning of the conflict, and in some cases there are two dotted lines, marking the first 3 weeks of the conflict.

An unusual trend is shown in Google Maps traffic over the marked 3 week period, depicted in Figure 2b. The Google Maps traffic rate increased by over 200% in the first days of the war, due to the wave of refugees and domestic transport, and after three weeks was only 25% higher than normal, gradually returning to normal over the following two months.

Another aspect that indicates the flow of refugees during these 2-3 weeks is the nature of mobile network usage. Based on Statcounter data we computed the ratio between mobile device traffic and desktop traffic. Figure 2c illustrates the mobile-to-desktop ratio in Ukraine, indicating a steep increase of over 10% in mobile device usage in the 2-3 weeks after the war started, and a large decrease in the ratio afterwards, which can be explained by the traveling and displacement during the first 2-3 weeks, which resulted in high mobile device usage compared to desktops, followed by a general decrease in mobile usage in the period that followed.

Figure 2d depicts the number of RIPE Atlas [16] active (connected) probes in Ukraine. Specifically, 222 probes were active in Ukraine on 23 February 2022, with an average of 219.9 probes during the year beforehand. The figure shows a steep decline to just 183 probes during the first two weeks of the war, with no significant changes afterwards. This behavior is not likely to be due to connectivity issues, which would likely result in more fluctuations. Instead, this behavior once again indicates a high correlation to the internal displacement and the refugee flow (Figure 2a).

Another aspect that was analyzed is the traffic rate in Ukraine at the beginning of the war and slightly beforehand. As shown in Figure 3, the traffic rate dropped during the first 2-3 weeks of the war, and then gradually started to increase.

We first observe the normalized daily peak traffic rate in Ukraine, measured by Cloudflare, as shown in Figure 3a. We found that the peak traffic rate in the two weeks after the war began was about 25% lower than during the two weeks beforehand. A similar trend is shown in Google’s data, depicted in Figure 3b; the normalized daily peak web search traffic rate decreased by 30% during the same period of time.

These low traffic rates lasted for 2-3 weeks and then gradually climbed back to the normal rates, a trend that can be explained by the refugee crisis and the internal population displacement, having a temporary effect that lasted until the displaced population managed to settle in.

An analysis of the types of mobile devices used in Ukraine during the first few weeks of the war reveals a steep change in the mobile device vendors used in Ukraine. The mobile vendor distribution in previous years, depicted in Figure 4b, focuses on four of the most common mobile device vendors during this period, and shows the gradual decrease in Nokia mobile devices, from 13% in 2015 to under 1% in 2021. However, during the first weeks of the war in 2022, as shown in Figure 4a, the rate of Nokia device usage increased back to 13%. This surprising change is explained by the increasing demand for mobile devices, causing people to revive older
and unused mobile devices. This trend gradually decreased during the first two months of the war.

![Mobile vendor usage in Ukraine](statcounter)

**Figure 4: Mobile vendor usage in Ukraine [Statcounter]**

### 3.2 Performance around Ukraine

One of the aspects we analyzed is the impact of the crisis on the Internet performance of countries around Ukraine. Since Poland was the country that received the largest number of refugees (Figure 1b), we take Poland as a case study, and focus our analysis in the current subsection on Poland.

We start by analyzing Internet access performance, based on Speedtest data, which shows the rank of each country for fixed networks and for mobile networks. Based on these measurements, we found that following the beginning of the war, four of the seven countries surrounding Ukraine were ranked lower than before the beginning of the war. Specifically, Poland was ranked lower in March 2022 than in February, both in fixed (one place lower) and in mobile tests (two places lower).

![Speedtest performance rank change in March 2022](statcounter)

**Figure 5: Performance impact on neighboring countries**

The traffic rate in Poland is another interesting metric in our analysis. We found that the Cloudflare traffic rate, shown in Figure 5b, increased by over 40% in the two weeks after the war started compared to the two weeks beforehand.

The mobile-to-desktop ratio in Poland, depicted in Figure 6a, demonstrates a significant increase in mobile usage in Poland shortly after the war started, which can be explained by the large number of people who were in transit and were mostly using mobile devices. This assessment is further confirmed by analyzing the mobile vendor usage, as shown in Figure 6b. As previously shown above (Figure 4), the most commonly used mobile vendor in Ukraine is Xiaomi. As illustrated in Figure 6b, the usage rate of Xiaomi devices has significantly increased in Poland during the first 2-3 weeks of the war, and then remained stable.

The mobile usage in Poland, as shown in Figure 6 suggests that the Ukrainian presence in Poland significantly increased during the first 3 weeks of the war, but then remained roughly stable afterwards despite the continued influx (Figure 1a), which confirms the fact that many of the refugees entering Poland continued onward to other countries.

![Mobile to desktop ratio in Europe](statcounter)

**Figure 6: Impact on mobile usage in Poland [Statcounter]**

![Mobile vendor usage](statcounter)

**Figure 7: Mobile to desktop ratio in Europe. The dotted line marks the beginning of the removal of COVID-19 travel restrictions.**

These results show that measurement data about mobile device usage provides a clear indication about refugee presence in Poland, which received the largest number of refugees.
However, this approach was not necessarily productive when we analyzed other countries in Europe. One of the important trends we noticed is the fact that starting from mid-January the mobile-to-desktop ratio has increased throughout Europe, following the removal of the COVID-19 travel restrictions [8] that were applied during the Omicron wave. With the travel restrictions removed, mobile traffic increased significantly with increasingly growing numbers of people traveling and using mobile devices. The high mobile network usage (Figure 6) in Poland was clearly due to the refugee wave, indicated by the date in which the steep increase started. However, when we tried to analyze mobile device usage in the rest of Europe, we found that it was deeply affected by the removal of the COVID restrictions, as shown in Figure 7, also causing significant change in mobile vendor usage.

The Poland case study shows that the refugee crisis affected not only the performance in Ukraine, but also in other countries. However, it should be noted that the indications we presented in the current subsection were less dominant in other countries we analyzed, and therefore these metrics would not be a practical means for mapping the refugees’ presence. However, the request rate to Ukrainian websites was found to be an effective metric for this purpose, as further discussed below.

4 MAPPING THE REFUGEE CRISIS

Many of the refugees who crossed the border to neighboring countries continued their journey to other countries. While the UNHCR keeps track of the number of refugees that have reached each of the neighboring countries, there is no accurate record of the number of refugees who stayed in these countries, or generally the number of refugees currently staying in each country in the world. One of the keys to understanding the crisis and helping the refugees is being able to assess how many refugees are staying in each country.

4.1 A Geographical Perspective

One of the main approaches we analyzed in order to map the Ukrainian presence in the world is to observe the visit rate to top Ukrainian websites from locations around the world.

We crossed information from two sources: the top Ukrainian sites and the number of visits per month were extracted from Similarweb [17], and the per-country visit rates for these websites are based on data from Cloudflare [7].

The visit rate to the top five Ukrainian websites is illustrated in Figure 8, focusing on the five countries in Europe which had the highest visit rates (excluding Ukraine). For example, 4.84% of the visits to google.com.ua came from Germany.

In our analysis we used data from 15 sites in order to estimate the Ukrainian presence in each country. These sites were taken from the list of top Ukrainian sites [17], eliminating international sites such as facebook.com and yandex.ru, and focusing on sites that are predominantly accessed by Ukrainian users.

We now present a brief overview of the model and assumptions that were used in our estimation method. We denote the total number of countries by \( N \), and the total number of Ukrainian website users, spread throughout the world, is denoted by \( S \). The number of individuals in each country is denoted by \( n_1, n_2, \ldots, n_N \), where the index indicates the country. We assume that every website visit has an equal probability of being accessed by each individual, and therefore the probability of being accessed from country \( i \) is \( p_i = n_i/S \), and thus we are analyzing a multinomial distribution. We would like to estimate the values \( p_i \) for \( i = 1, \ldots, N \), representing the proportional part of the population in each country.

For a given set of measurements which includes \( A \) website visits, with \( x_1, \ldots, x_N \) specifying the number of visits from each country, we can evaluate the estimated values \( \hat{p}_i \) by using a Maximum Likelihood (ML) estimator for multinomial distribution (e.g., [3]), which is given by \( \hat{p}_i = x_i/A \).

We analyzed \( K \) websites (\( K = 15 \) in our analysis), where for each site \( j \) the number of visits per month, \( A_j \), was extracted from [17]. For each site \( j \) we used the values \( x_{i,j} \) from [7], specifying the relative number of visits from country \( i \) to site \( j \), and thus \( x_{i,j} \cdot A_j \) is the absolute number of visits per month. Hence, for each country \( i \) we have \( x_i = \sum_{j=1}^{K} x_{i,j} \cdot A_j \). Thus, the ML estimator is as follows:
The results of the estimation are illustrated in Figure 9, showing for each of the top countries the estimated number of Ukrainians, expressed as the percentage from the world’s Ukrainian population. For example, it is estimated that 2.03% of the Ukrainian population is located in Poland, which is about 1 million people. The histogram presents the estimated percentage, computed based on the ML estimator of Eq. 1, as well as the standard deviation (STDEV) compared to the estimated value for each country. We note that the STDEV is high for Russia, indicating a low confidence level. However, the estimated presence in Russia is high, indicating over 1 million people, suggesting that the number reported by UNHCR, 0.77 million (Figure 1b), possibly does not reflect the full picture, and potentially confirms the Russian announcement about transferring over 1 million Ukrainians to Russia.

Low STDEV values were computed for Germany and Poland and other European countries, showing a higher confidence level in our results.

This analysis can be used as complementary data to the daily UNHCR data of Ukrainian border crossing. Notably, our analysis indicates that there is significant presence of Ukrainians in the US, Germany, France, the Czech Rep. and the Netherlands, as well as other European countries, and not only in Ukraine’s neighboring countries.

Focusing on google.com.ua as an example of one of the top websites, Figure 10 illustrates a detailed comparison of the visit rate in various countries in Europe.

4.2 A Time Domain Perspective

While Ukrainian website visit rates can provide a useful insight into Ukrainian presence in countries throughout the world, these figures are also affected by the Ukrainian diaspora prior to the war. According to [21] the number of Ukrainian citizens spread throughout the world in June 2020 was approximately 6.1 million.

Notably, historic data about website visits can be used to assess the trends in the geographic location of Ukrainian website visits, and thus can assist in a time domain analysis of the trends in the flow of refugees.

Our analysis focused on historic data that is publicly available; we found that Statcounter data includes search engine host market share on a per-country basis, and specifically includes per-country data about the visits to google.com.ua. Notably, the data is expressed as the relative visit rate compared to local search engine requests. For example, 0.07% of the search engine requests in Germany went to google.com.ua in May 2022. Conversely, the data on Figures 8 and 9 was expressed as the percentage from the total visits in the analyzed sites. Thus, in order to analyze the historic data, we normalize the data from each country, as shown in Figure 11, illustrating the trends of the Ukrainian presence in each of these countries in a relative manner. For example, Figure 11b shows the vast influx of refugees into Poland on the first two weeks of the war, while the influx into Germany reached its peak on the third week of the war, as shown in Figure 11c.

Combining the two types of analytics provides a broader picture: website visit analytics (Figures 8 and 9) provide the geographic distribution of Ukrainian presence, and historic data (Figure 11) indicates the trends of the refugee flow.

5 RELATED WORK

To the best of our knowledge, we are the first to publish academic research connecting Internet measurements to the Ukrainian refugee crisis. Several blogs that are related to the conflict in Ukraine have been published, among them [2] surveyed the status of the Internet infrastructure at the beginning of the conflict, and [1] studied the resilience of the Internet during the conflict.
The COVID-19 pandemic has widely affected the usage of the Internet over the last few years. The studies published on this topic are an example of how Internet traffic might be used to analyze a worldwide crisis. Specifically, [6] studied the impact of the COVID-19 pandemic on the Internet latency. [20] studied the changes of Internet traffic in on-campus dormitories during the pandemic lockdown, providing a focused lens on the behavior of undergraduate student population during the pandemic. In [10], the authors examined the effect of government lockdowns on Internet traffic, finding a significant increase of 15-20% of traffic volume. [5] provided a perspective of the scale of Internet traffic growth and how well the Internet coped with the increased demand as seen from Facebook’s edge network during the COVID-19 pandemic. In [13], the impact of the COVID-19 crisis on a UK mobile network operator is studied. [9] characterized how Internet traffic and application demands change over a year in lockdown.

6 CONCLUSION

In this paper we analyzed how Internet measurement data can be used to map the Ukrainian refugee crisis. We demonstrated the unique footprint that the refugee crisis had on the Internet performance in Ukraine and around it, and introduced a novel approach to estimate the refugee distribution using publicly available information about website visits. Notably, the proposed methodology can achieve higher accuracy if more detailed data will become publicly available. Specifically, detailed historic data about website visit rates can provide a much deeper insight not only into the distribution of individuals throughout the world, but also about the trends and the flow of refugees. We hope that this work can contribute to the ongoing effort to support the refugees.

REFERENCES

[1] E. Aben. The Resilience of the Internet in Ukraine. Ripe labs, 10 March 2022, https://labs.ripe.net/author/emileaben/the-resilience-of-the-internet-in-ukraine/, 2022.
[2] E. Aben. The Ukrainian Internet. Ripe labs, 28 February 2022, https://labs.ripe.net/author/emileaben/the-ukrainian-internet/, 2022.
[3] J. Arnold. Maximum Likelihood for the Multinomial Distribution. https://blog.jakuba.net/maximum-likelihood-for-multinomial-distribution/, 2018.
[4] BBC. Ukraine war: Putin being misled by fearful advisers, US says. Bbc news, 31 March 2022, https://www.bbc.com/news/world-europe-60936117, 2022.
[5] T. Böttger, G. Ibrahim, and B. Vallis. How the internet reacted to covid-19: A perspective from facebook’s edge network. In Proceedings of the ACM Internet Measurement Conference, pages 34–41, 2020.
[6] M. Candela, V. Luconi, and A. Vecchio. Impact of the covid-19 pandemic on the internet latency. A large-scale study. Computer Networks, 182:107495, 2020.
[7] Cloudflare. Cloudflare Radar. https://radar.cloudflare.com/, 2022.
[8] DW. Countries are relaxing restrictions after omicron spikes. 17 January 2022, https://www.dw.com/en/countries-are-relaxing-restrictions-after-omicron-spikes/a-60450812, 2022.
[9] A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poese, C. Dietzel, D. Wagner, M. Wichthuber, J. Tapiador, N. Vallina-Rodriguez, et al. A year in lockdown: how the waves of covid-19 impact internet traffic. *Communications of the ACM*, 64(7):101–108, 2021.

[10] A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poese, C. Dietzel, D. Wagner, M. Wichthuber, J. Tapiador, N. Vallina-Rodriguez, O. Hohlfeld, and G. Smaragdakis. The lockdown effect: Implications of the COVID-19 pandemic on internet traffic. In *IMC ’20: ACM Internet Measurement Conference*, Virtual Event, USA, October 27-29, 2020, 2020.

[11] Google. Google Transparency Report. https://transparencyreport.google.com/traffic/overview, 2022.

[12] L. Kelly. More than 1 mln people evacuated from Ukraine to Russia since Feb. 24, says Lavrov. Reuters, 30 april 2022, https://www.reuters.com/world/more-than-1-mln-people-evacuated-ukraine-russia-since-feb-24-says-lavrov-2022-04-30/, 2022.

[13] A. Lutu, D. Perino, M. Bagnulo, E. Frias-Martinez, and J. Khangosstar. A characterization of the covid-19 pandemic impact on a mobile network operator traffic. In *Proceedings of the ACM internet measurement conference*, pages 19–33, 2020.

[14] T. Mizrahi and J. Yallouz. Internet Performance in the 2022 Conflict in Ukraine: An Asymmetric Analysis. Arxiv, 2022.

[15] A. Pita and R. S. Costa. Ukrainian exodus could be Europe’s biggest refugee crisis since World War II. El pais, 3 march 2022, https://english.elpais.com/international/2022-03-03/ukrainian-exodus-could-be-europes-biggest-refugee-crisis-since-world-war-ii.html, 2022.

[16] RIPE NCC. Ripe atlas. https://atlas.ripe.net/, 2022.

[17] Similarweb. Top Websites Ranking. https://www.similarweb.com/top-websites/ukraine/, 2022.

[18] Speedtest. Speedtest global index. https://www.speedtest.net/global-index, April, 2022.

[19] Statcounter. Global stats. https://gs.statcounter.com/, April, 2022.

[20] A. Ukani, A. Mirian, and A. C. Snoeren. Locked-in during lock-down: undergraduate life on the internet in a pandemic. In *Proceedings of the 21st ACM Internet Measurement Conference*, pages 480–486, 2021.

[21] UNHCR. Ukraine refugee situation. https://data2.unhcr.org/en/situations/ukraine, April, 2022.

[22] Worldometer. Population. https://www.worldometers.info/, 2022.