Characterizing and Modeling Traffic of Communication and Collaboration Apps Bloomed With COVID-19 Outbreak

Idio Guarino, Giuseppe Aceto, Domenico Ciuonzo, Antonio Montieri, Valerio Persico, Antonio Pescapè
University of Napoli "Federico II" (Italy)
{idio.guarino, giuseppe.aceto, domenico.ciuonzo, antonio.montieri, valerio.persico, pescape}@unina.it

***Abstract*—In this work, we address the characterization and modeling of the network traffic generated by communication and collaboration apps which have been the object of recent traffic surge due to the COVID-19 pandemic spread. In detail, focusing on five of the top popular mobile apps (collected via the MIRAGE architecture) used for working/studying during the pandemic time frame, we provide characterization at trace and flow level, and modeling by means of Multimodal Markov Chains for both apps and related activities. The results highlight interesting peculiarities related to both the running applications and the specific activities performed. The outcome of this analysis constitutes the stepping stone toward a number of tasks related to network management and traffic analysis, such as identification/classification and prediction, and modern IT management in general.**

***Index Terms*—communication apps; collaboration apps; COVID-19; encrypted traffic; Markov models; traffic characterization; traffic modeling;**

I. Introduction

Due to the COVID-19 pandemic, many governments have imposed lockdown periods that have forced millions of citizens to spend more time at home and adopt "smart working" modes wherever possible. The implementation of these restrictions has generated increased demand in terms of Internet traffic from residential users [1] for remote working, entertainment, commerce, and education. In addition, the sudden change imposed on the lifestyles of millions of people has generated abrupt and noticeable changes to the nature of the traffic flowing through the Internet. Recent studies show that these events have also had a non-negligible impact on the performance of networks, which, having to cope with a suddenly higher demand, have guaranteed lower performance levels compared to the period immediately before (e.g., showing increases in delay variability and loss rates) [2]. As networks have to respond to a greater and qualitatively different demand, new tools are needed for the effective and efficient management of traffic generated by users when accessing remote services. This situation therefore introduces new and interesting research challenges, aimed at identifying innovative solutions that support the monitoring, management, and engineering activities of the networks themselves, in order to guarantee their correct functioning while respecting the expected security levels.

In view of these considerations, the contributions of this work can be summarized as follows:

- we investigate reports of mobile apps' usage increased during COVID-19 lockdown, and select five mobile apps among the most used in the category *communication and collaboration*;
- we collect and reliably label network traffic generated by the selected mobile apps by means of a dedicated architecture (MIRAGE);
- we characterize the collected traffic in terms of bitrate, packet-rate, and upstream/downstream ratio (in bit volume and packet volume), at different granularities, namely at *trace* and *flow* level, for both *apps* and related *activities*;
- we model the sequence of packet sizes, inter-arrival times, and directions with *Multimodal Markov Chains*, at *per-activity* granularity.

The resulting analysis highlights similarities and differences among the apps and the activities, of interest for app classification and identification, network and service planning and management, and for IT managers in general.

The paper is organized as follows. Section II surveys related studies characterizing and modeling encrypted traffic during/after the pandemic phase, positioning our work against related literature. Section III describes the considered capture and ground-truth generation architecture, as well as the dataset collected. The experimental analysis is provided in Sec. IV. Finally, Sec. V provides conclusions and future perspectives.

II. Related Work

Following the spread of the COVID-19 pandemic on a global scale, several works have analyzed its impact on the Internet. In this regard, Feldmann et al. [1] show that during the March-June 2020 lockdown period, the volume of European Internet traffic increased by up to $+20\%$. This increase is mainly due to residential traffic generated by applications for social interaction and smart-working (up to $+200\%$), while there was a reduction in traffic from educational networks (-55%). Complementary results are reported by Lutu et al. [3], which analyze trends in mobile traffic, highlighting, in

This work was carried out by Dr. Idio Guarino under the grant "O. Carlini" funded by GARR, the Italian national network of University and Research.

addition to a predictable decrease in user mobility (-50%), a substantial increase in voice traffic ($+120\%$) that impacted network performance, with evidence of packet loss, albeit for a limited period of time. Further degradation of network services was seen in regions with less developed infrastructures, while North American and European networks returned to pre-COVID levels of operation relatively quickly [4]. As far as Italy is concerned, Favale et al. [5] show how the increased use of digital tools such as collaboration platforms, VPNs and remote desktop services has impacted the university network of the Politecnico di Torino, which reached traffic peaks of $1.5\ \mathrm{Gbit/s}$ during March/April 2020, while still managing to guarantee the operation of services. Nevertheless, Candela et al. [2] show that in Italy there has been an increase in both latency variability and packet loss rates, thus exacerbating concerns about the digital divide. Finally, Affinito et al. [6] analyze the impact on the use of different categories of Internet applications (i.e. *Video*, *SocialMedia*, *Messaging*, and *Collaboration Tool*), by analyzing websites and domains used during the enforcement of the lockdown and other social distancing measures. They show that during the lockdown period, the most used applications were those for online video content (i.e. `Youtube` and `Netflix`) and social interaction (i.e. `Facebook`, `Whatsapp`, and `Skype`). In particular, they also highlight the increased usage of `Zoom` in the second half of March 2020, in conjunction with the beginning of the adoption of the tool for smart-working and e-teaching.

With respect to the aforementioned work, in this paper we focus on characterizing the traffic of communication and collaboration apps that experienced the most increase in usage, with significant concordance with Affinito et al. [6] (4 out of 5 of the domain names they consider in the "collaboration tool" group correspond to the apps we analyze). The other works focus on different traffic characteristics, namely volumes [1, 3, 5] and delays [2, 5], in a diverse set of network infrastructures. As our analysis employs a dataset we collected from the campus network, our networking scenario is more similar to the one in Favale et al. [5], with which we have also one analyzed application in common (namely, `Teams`).

From the methodological viewpoint, the closest related work is Aceto et al. [7], our previous contribution focusing on mobile video traffic modeling and prediction. With respect to that work, we here focus on characterization (rather than prediction), and consider a different set of apps (selected according to their sudden increase in popularity associated to lockdown). Moreover, we deepen the analysis at *per-activity* granularity (a level of detail unavailable in the previous work), characterizing the traffic for different usages of each app (namely, *webinar, video-call, and video-conference*).

III. Dataset

Herein we describe the dataset leveraged to conduct the analyses, briefly detailing also the architecture utilized for its collection in Sec. III-A. Apps' selection rationale and details on activities performed are then given in Sec. III-B.

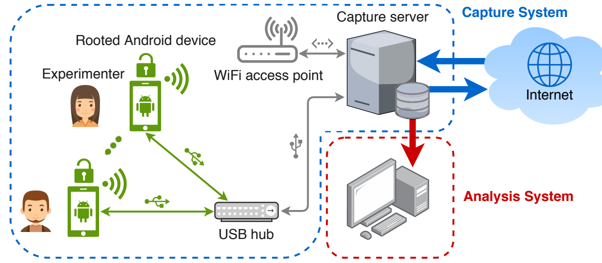


Figure 1: Bird's eye view of MIRAGE architecture [8].

A. Traffic Collection and Ground Truth Generation

The dataset was collected leveraging the MIRAGE architecture [8] (conveniently optimized to capture video traffic) in the ARCLAB laboratory at the University of Napoli "Federico II", using three mobile devices: a Google Nexus 6 (Android 10) and two Samsung Galaxy A5 (Android 6.0.1). MIRAGE architecture is depicted in Fig. 1 and consists of two main components: the *Capture System* and the *Analysis System*. The Capture System provides connectivity (via a WiFi access point) to *rooted mobile devices* that generate the traffic when human *experimenters* utilize the video apps, and sends/receives commands/responses (with the Android Debug Bridge via an USB hub) on an off-band channel. The simultaneous capture of multiple devices is handled via their MAC addresses.

The dataset has been constructed by involving students and researchers[1] from the University of Napoli "Federico II" who took on the role of experimenters. Specifically, in each capture session the experimenter performed a specific activity, so as to obtain a traffic dataset that reflects the common usage of considered apps (see Tab. I).[2] The duration of each capture session (spanning from 15 to 80 minutes) depends on the type of activity carried out using the specific application, each resulting in a PCAP traffic trace and additional system log-files with ground-truth information.

Such log-files are exploited to reliably label each biflow with the corresponding *Android package-name* that matches the 5-tuple by considering established network-connections via `netstat`. To this end, the Analysis System extracts the Android package to which the socket belongs, namely which is listening on the *<IP:port>* pair of the socket.

B. Apps' selection rationale

Conference and collaboration tools have experienced a huge increment of their utilization when "stay-at-home" orders were issued worldwide. Indeed, during this time-frame, they have been used to conduct both online meetings in businesses and classes in schools and universities, in addition to maintain social interaction. Herein we focus on **five** conference and collaboration apps, reported in Tab. I: `GotoMeeting`, `Skype`,

[1]We highlight that the captures were carried out by adhering to the distancing/mask-wearing rules prescribed by regional/national decrees in force at the moment of the collection.

[2]Each traffic capture session has been performed with the up-to-date version of the app. Also, to limit background traffic, network access has been disabled for all the apps but the one under test.

Table I: Communication and collaboration apps considered. Performed activities, number of packets (#pkts), total duration (Dur), and traffic volume (Vol) are reported for each app.

App	Webi	VCall	VConf	#pkts [M]	Dur [H:m]	Vol [GB]
GotoMeeting	✓			7.32	9:54	1.94
Skype	✓	✓	✓	2.21	3:55	1.41
Teams	✓	✓	✓	10.27	14:01	5.65
Webex	✓	✓	✓	2.60	3:30	1.84
Zoom	✓	✓	✓	5.46	6:31	4.26

Teams, Webex, and Zoom. These apps have been selected considering both popularity and utilization increment.

Specifically, according to the latest Sandvine report [9], Zoom has obtained the steepest increment with its traffic scaling to orders of magnitude, followed by Webex, GotoMeeting, Teams, BlueJeans (whose traffic we are currently collecting), and Skype. More specifically, the App Annie's market analysis [10] reports that during 15th–21st March 2020, Zoom was downloaded (from the Google Play Store) $14\times$ more than the weekly average during Q4 2019 in the US, $20\times$ more in the UK, and impressively $55\times$ more in Italy. Similarly, Teams also experienced significant growth in France and Italy with $16\times$ and $30\times$ more downloads than the Q4 2019 weekly average, respectively. Last but not least, the considered apps have been extensively exploited for remote teaching in Italian [11] (employing the national GARR network) and European [12] institutions and universities.

Table I summarizes the mobile apps used in this study, highlighting also the activities carried out with each, and the amount of traffic collected in terms of packets, duration, and volume. Specifically, the activities performed are:

- *Webinar (Webi)*: a live event involving many attendees and one presenter who transmits his/her own audio together with slides and/or his/her own video (e.g., seminar or online lesson).
- *Video-call (VCall)*: a live event involving just two participants who transmit both audio and video traffic.
- *Video-conference (VConf)*: a live event involving more than two participants broadcasting audio/video traffic.

IV. Experimental Analysis

In this section, we provide a characterization analysis—both at trace (Sec. IV-A) and flow level (Sec. IV-B)—and modeling evaluation (Sec. IV-C) of the considered apps/activities. The provided analysis can be useful for different network-related tasks supporting the detailed understanding of the network traffic, which is critical for properly managing the network, as well as for identifying peculiarities (e.g., network fingerprints) of both apps and specific activities which can be leveraged for modeling purposes to support traffic classification and identification tasks.

A. Per-trace characterization

Hereinafter we perform a characterization of the collected traffic in terms of (a) bitrate and (b) packet-rate. In detail, we refer to *aggregated rates*, i.e. computed by considering the whole traffic either generated or received by the app (i.e. including concurrent bidirectional flows) during a traffic capture. In the following, aggregate metrics are computed considering a (non-overlapping) window size $\Delta = 5\,\mathrm{s}$.[3] Formally, for a capture starting at time t_0 and having duration D, the i^{th} aggregation interval gathers all packets whose arrival time falls within $[t_0 + (i-1)\Delta, t_0 + i\Delta)$, with $i \in \{1, 2, \cdots \lceil \frac{D}{\Delta} \rceil\}$. Figures 2a–2b and 2d–2e report the boxplots of bitrate and packet-rate (both computed over 5-second intervals), respectively, considering upstream and downstream directions separately. For each app, the distribution is broken down across the specific activities highlighted with different colors.

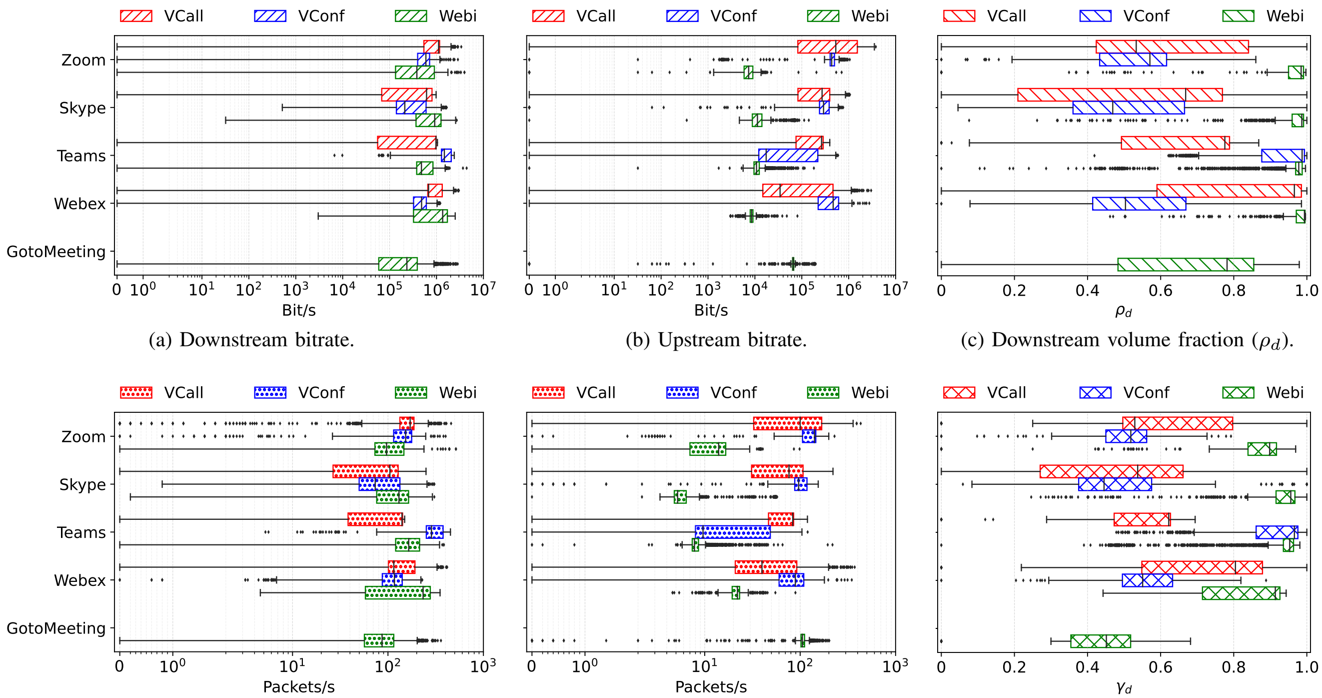
Considering the **downstream bitrate** (Fig. 2a), all the apps and activities result in a median value between $200\,\mathrm{Kbps}$ and $1.5\,\mathrm{Mbps}$, with the highest values up to $4\,\mathrm{Mbps}$. Neither apps nor activities can be straightforwardly ranked according to their downstream bitrate. Variability depends on the specific app and activity, with IQRs varying from $282\,\mathrm{Kbps}$ to $1.38\,\mathrm{Mbps}$, and *VCall* bitrate always exposing higher variability (in terms of IQR) than *VConf*. All the apps but Teams highlight higher downstream bitrate (on median) for *VCall* traffic compared to *VConf*. This suggests different strategies enforced when dealing with multiple parallel media streams which are likely delivered at lower quality.

Moving to the **upstream bitrate** (Fig. 2b), observed values span over a larger interval, with lowest median values as low as less than $10\,\mathrm{Kbps}$. Interestingly, Zoom and Skype do not show differences between *VCall* and *VConf* in terms of median values, in spite of a higher variability observed for *VCall*. On the other hand, Teams and Webex traffic results in significant discrepancies when comparing *VCall* and *VConf* upstream bitrate. For all the apps, *Webi* activity is related to the lowest bitrate observed, with values typically lying around $10\,\mathrm{Kbps}$ (with the exception of GotoMeeting, having a median at 60–$70\,\mathrm{Kbps}$). Such an outcome is in line with the commonly fewer interactions performed in the case of *Webi* activity.

Concerning **downstream** and **upstream packet-rate** (Figs. 2d and 2e, respectively), all the apps receive traffic (i.e. downstream direction) at 80–$300\,\mathrm{packets/s}$, on median. Looking at packets in the opposite direction, their rate ranges—on median—from few packets per second ($\approx 5\,\mathrm{packets/s}$) up to $\approx 140\,\mathrm{packets/s}$. More specifically, the upstream packet-rate seems to be strongly correlated to the specific activity (with *Webi* always exposing the lowest rate). Remarkably, while no major discrepancy can be spotted when considering *VCall* and *VConf* activities for the same app, the *VConf* upstream traffic on Teams is almost one order of magnitude lower, than the *VCall* traffic generated by the same app, on median.

Finally, considering **traffic direction**, we quantify the fraction of downstream volume and packet by means of ρ_d and γ_d, which are defined as follows: $\rho_d = \frac{B_d}{B_d + B_u}$ and $\gamma_d = \frac{P_d}{P_d + P_u}$, where B_d and B_u refer to the number of downstream and upstream bytes, respectively, and P_d and P_u refer to the

[3]Further significant values of $\Delta = 10, 30, 60$ s have been tested without showing notable discrepancies.



(a) Downstream bitrate. (b) Upstream bitrate. (c) Downstream volume fraction (ρ_d).

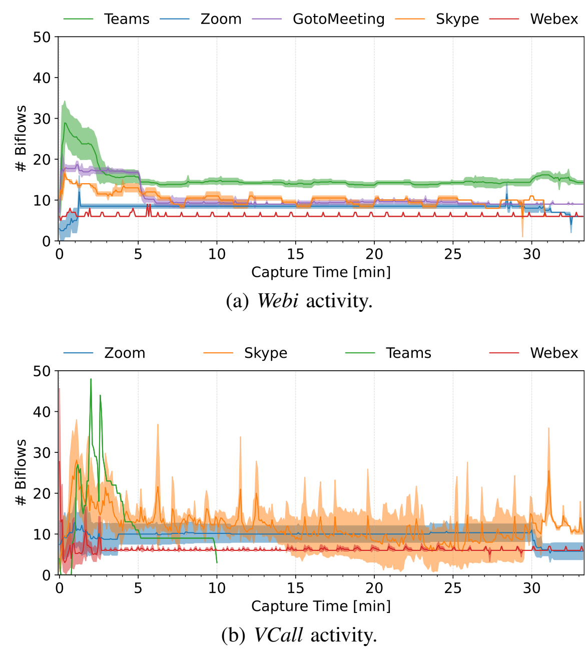
(d) Downstream packet-rate. (e) Upstream packet-rate. (f) Downstream packet fraction (γ_d).

Figure 2: Downstream bitrate (a), upstream bitrate (b), percentage of downstream bitrate (c), downstream packet-rate (d) upstream packet-rate (e), and percentage of downstream packets (f). Values are evaluated over time intervals of $\Delta = 5$ s. Boxes report the 1^{st} and 3^{rd} quartiles (1Q and 3Q, respectively), while whiskers mark 1Q$-$1.5 IQR and 3Q+1.5 IQR, where IQR=3Q$-$1Q. Black diamonds highlight outliers.

number of downstream and upstream packets, respectively. Accordingly, Figs. 2c and 2f report the distribution of these two metrics computed over 5-second intervals. Focusing on *VCall* and *VConf* activities, median values witness that the usage of `Zoom` results in more balanced traffic in terms of direction, with $\rho_d \approx 0.5$. On the other hand, for both `Teams` and `Webex` the median of the distribution of ρ_d sits on remarkably higher values in case of *VCall* activity. Interestingly, this consideration does not hold also for *VConf*. More in general, the latter activity results in a more stable behavior (less variability can be spotted when looking at IQR) with respect to *VCall*. The above considerations about the resulting distribution also apply to γ_d, suggesting the presence of almost-constant-sized packets. Finally, traffic is remarkably unbalanced towards the downstream direction when considering *Webi* traffic, with ρ_d consistently ≈ 1 for all the apps but `GotoMeeting` (whose median $\rho_d \approx 0.8$). Still considering *Webi*, it is worth noticing that γ_d spans over slightly-lower values, witnessing the non-negligible presence of small upstream packets, likely due to signaling operations.

B. Per-flow characterization

The aggregated behavior previously discussed is the result of the mixture of multiple concurrent traffic flows between the client running the communication and collaboration app and a variety of servers. The amount and the characteristics of these flows depends on the app and the activities.



(a) *Webi* activity.

(b) *VCall* activity.

Figure 3: Amount of concurrent biflows (*mean* $\pm$ *std* across different captures) in each 5-second slot.

Considering two exemplifying activities (i.e. *Webi* and *VCall*), Fig. 3 shows the **amount of concurrent bidirectional flows** (viz. biflows) considering non-overlapping 5-second

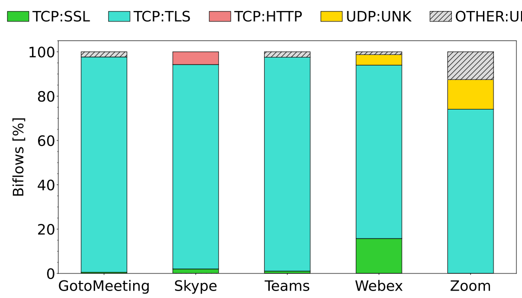


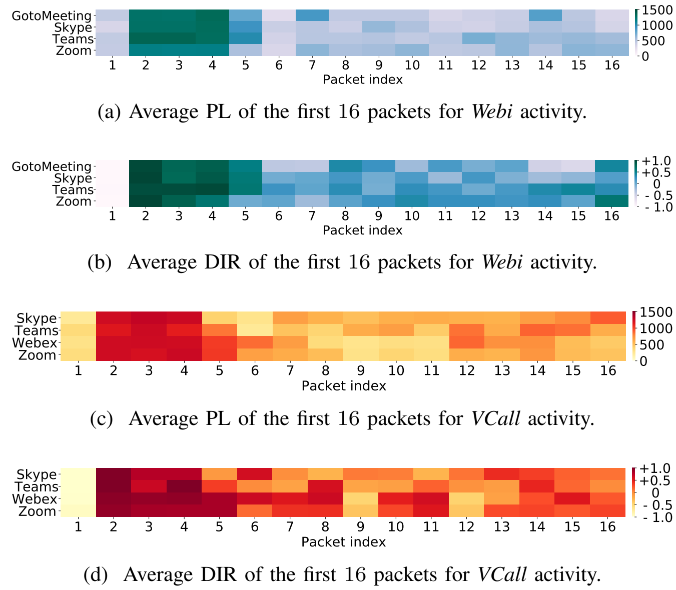
Figure 4: Protocol distribution in terms of biflows. *UNK* stands for *unknown*, *SSL* stands for *undetected* version of *SSL/TLS*.

windows as described in the previous section. The figure reports the breakdown across the different apps by showing the mean and the standard deviation of the (open) biflow count in different captures. For both activities, a high number of biflows is typically opened by the client at the very beginning of the activity. These concurrent communications are possibly related to authorization/accounting services, which are expected to be located on dedicated hosts. Then, the number of open flows settles to a lower value (which varies with both the activity and the app and ranges from ≈ 6 to 15 biflows). While this plateau value appears to be more stable for *Webi*, more evident discrepancies across different captures are observed for *VCall* (especially for `Skype`). Remarkably the above pattern does not apply to `Zoom`, which at the beginning of both activities incrementally opens concurrent flows up to the plateau value.

Figure 4 reports the **traffic composition in terms of the adopted protocols** as obtained by leveraging the `tshark` dissector. The reported breakdown highlights the remarkable presence of SSL and TLS, which together account for a large fraction of the biflows for all the apps (from 56% for `Zoom` to 83% for `GotoMeeting`). Also, both `Webex` and `Zoom` show a non-negligible part of UDP traffic, which is possibly adopted for delivering media streams without incurring the issues of TCP closed loop. Finally, a minor fraction of `Skype` traffic is delivered through (cleartext) HTTP.

Conversely, by zooming on the amount of information exchanged across these flows, Fig. 5 reports the **sequence of payload length and direction** of app-level packets (packets with no payload are discarded since they reflect transport-layer signaling which does not depend on the nature of the app or the performed activity). In more detail, the analysis focuses on the first 16 data packets of the biflows resulting from *Webi* (Figs. 5a–5b) and *VCall* (Figs. 5c–5d) activities. Specifically, for a given app, the figure reports the average value on all biflows for each packet index.

Considering the payload length (PL), for both activities, all apps show a similar behavior when observing the first 5 packets: a small PL ($\approx$ 500 B) for the first packet followed by 4 packets with larger size ($\approx$ 1200 B, on average). Considering also the direction (DIR), the recurrent pattern (first packet in upstream direction followed by a number of packets in the opposite direction) reflects a typical client-



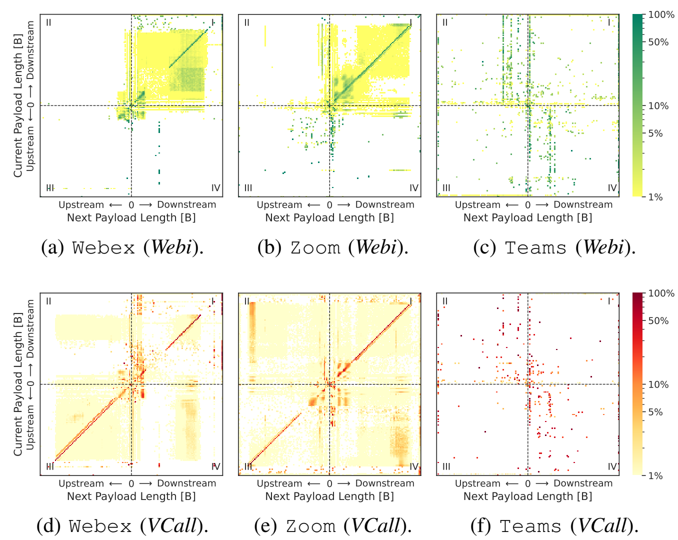
(a) Average PL of the first 16 packets for *Webi* activity.

(b) Average DIR of the first 16 packets for *Webi* activity.

(c) Average PL of the first 16 packets for *VCall* activity.

(d) Average DIR of the first 16 packets for *VCall* activity.

Figure 5: Properties of biflows time series with respect to PL and DIR: *Webi* (a, b) and *VCall* (c, d). The downstream and upstream packet direction was mapped on $+1$ and -1 values, respectively.



(a) `Webex` (*Webi*). (b) `Zoom` (*Webi*). (c) `Teams` (*Webi*).

(d) `Webex` (*VCall*). (e) `Zoom` (*VCall*). (f) `Teams` (*VCall*).

Figure 6: Transition matrices of payload length and packet direction for `Webex` (a, d), `Zoom` (b, e) and `Teams` (c, f), and related to *Webi* and *VCall* activities.

server interaction (e.g., TLS initial handshake). Finally, some app-dependent patterns are also evident, e.g., packets #7 and #14 for `GotoMeeting` (*Webi* activity) likely result to be in upstream direction and of large PL ($\approx$ 1000 B, on average).

C. Markov modeling of apps and activities

Herein we aim to bring out the peculiar characteristics of each app by modeling its traffic through (*Multimodal*) *Markov Chains*, following [7]. Specifically, for each app we consider *jointly* the (transport-layer) PL and the DIR of the packets of each biflow, and we derive the corresponding transition matrix, where $<(p_i, d_i), (p_j, d_j)>$ represents the probability that the next packet comes with a PL of p_j bytes and a direction d_j if the last observed packet has a PL of p_i bytes and a direction d_i. To this end, we performed two pre-processing operations: *(i)* removal of null-payload packets, that are assumed to be non-informative since they do not reflect the behaviors of the app but only the mechanisms of the transport layer; *(ii)* discretization of the PL through an unsupervised adaptive binning procedure based on K-means, leading to 80 bins.[4]

In Fig. 6, we show the transitions matrices obtained for, (a) `Webex`, (b) `Zoom`, and (c) `Teams` by considering (i) *Webi* and (ii) *VCall* activities separately. From visual inspection of results (higher values are associated to darker color), **different patterns** can be identified in the matrices depending on the app and the activity considered. Specifically, in both cases, for a given activity we can observe the presence of dark pattern on the main diagonal ($\nearrow$) and darker areas (blocks $\blacksquare$) common to `Webex` and `Zoom` where the former indicates the tendency of the two apps to generate pairs of packets with equal PL and DIR while the latter indicate that apps are more likely to generate packets that match values that fall within a similar set of values. Differently, in the case of `Teams` several vertical patterns ($\uparrow$) are visible in the II and IV quadrants, highlighting the presence of highly probable-values of downstream (resp. upstream) PLs within the observed traffic, namely independent of the current upstream (resp. downstream) value.

Furthermore, by comparing the transition matrices corresponding to different activities of the same app, we can observe different patterns, underlining the peculiarities of the specific activity. One relevant case is represented by `Webex` and `Zoom`, whose diagonal patterns ($\nearrow$) reflect the nature of the activity carried out by the experimenter. Specifically, considering the *Webi* activity, a diagonal pattern is obtained only in correspondence of the downstream direction: this can be explained since the experimenters were not active during the traffic capture, thus implying low upstream traffic. Conversely, in the case of *VCall*, the diagonal pattern appears for both directions: indeed, such activity involves the active (symmetric) participation of the users. Finally, *Webi* activity on `Teams` implies several vertical patterns ($\uparrow$), as opposed to the sparser case of *VCall*, associated to the 1^{st} (≈ 40B) and 30^{th} (≈ 500B) bin in both directions.

V. Conclusions and Future Directions

Network traffic composition underwent significant changes due to the COVID-19 outbreak which has remarkably boosted the popularity of C&C apps. Motivated by the fact that knowing the characteristics and the peculiarity of the traffic crossing the network is critical for a number of network-related task such as identification/classification and prediction, we addressed the characterization and modeling of the network traffic generated by five C&C apps (`GotoMeeting`, `Skype`, `Teams`, `Webex`, and `Zoom`) whose popularity suddenly increased with the COVID-19 outbreak. We collected ≈ 40 hours of network traffic in the context of MIRAGE project leveraging the volunteer experimenters. For each app, we considered three activities relevant for remote working/studying: *Webi*, *VCall*, and *VConf*.

First, our study provided a characterization at trace level of the traffic, identifying peculiar characteristics in terms of downstream and upstream bit rate and packet rate as well as downstream volume and packet fraction. Then, we inspected the traffic at flow level, highlighting how the presence of concurrent biflows depends upon both app and activity, and also follows interesting trends with the time. Also, we investigated traffic composition in terms of adopted protocols, and noticed the major presence of TLS and SSL biflows, save from some notable exceptions (`Zoom`, `Webex` and `Skype`). Finally, we provided a modeling analysis by means of Markov chains, highlighting some similar transition patterns for different apps performing the same activity (e.g. `Webex` and `Zoom` when performing both *VCall* and *Webi*) and different patterns for some apps (e.g. `Teams`).

Future directions of research will focus on capitalizing this knowledge for fine-grained prediction and classification of C&C apps (a) by Markov chains, (b) by adopting novel machine/deep learning models and (c) even via hybrid (but interpretable) combinations of both.

[4]Such choice was obtained by evaluating the quantization error of PL values versus the number of bins and by choosing the value balancing the error and the model complexity.

References

[1] A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poese, C. Dietzel, D. Wagner, M. Wichtlhuber, J. Tapiador, N. Vallina-Rodriguez, O. Hohlfeld, and G. Smaragdakis, "The lockdown effect: Implications of the COVID-19 pandemic on Internet traffic," in *ACM Internet Measurement Conference (IMC)*, 2020, p. 1–18.

[2] M. Candela, V. Luconi, and A. Vecchio, "Impact of the COVID-19 pandemic on the Internet latency: A large-scale study," *Computer Networks*, vol. 182, p. 107495, 2020.

[3] 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 *ACM Internet Measurement Conference (IMC)*, 2020, p. 19–33.

[4] T. Böttger, G. Ibrahim, and B. Vallis, "How the Internet reacted to COVID-19: A perspective from Facebook's edge network," in *ACM Internet Measurement Conference (IMC)*, 2020, p. 34–41.

[5] T. Favale, F. Soro, M. Trevisan, I. Drago, and M. Mellia, "Campus traffic and e-Learning during COVID-19 pandemic," *Computer Networks*, vol. 176, p. 107290, 2020.

[6] A. Affinito, A. Botta, and G. Ventre, "The impact of COVID on network utilization: an analysis on domain popularity," in *IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2020, pp. 1–6.

[7] G. Aceto, G. Bovenzi, D. Ciuonzo, A. Montieri, V. Persico, and A. Pescapé, "Characterization and prediction of mobile-app traffic using Markov modeling," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 907–925, 2021.

[8] G. Aceto, D. Ciuonzo, A. Montieri, V. Persico, and A. Pescapé, "MIRAGE: Mobile-app traffic capture and ground-truth creation," in *4th International Conference on Computing, Communications and Security (ICCCS)*, 2019, pp. 1–8.

[9] Sandvine, "The Global Internet Phenomena Report COVID-19 Spotlight." May 2020.

[10] Lexi Sydow - App Annie, "Video Conferencing Apps Surge from Coronavirus Impact." March 2020.

[11] Fondazione CRUI, "COVID-19 | Strumenti per la didattica digitale (in italian)." June 2021.

[12] European University Institute, "Software available at the EUI." June 2021.