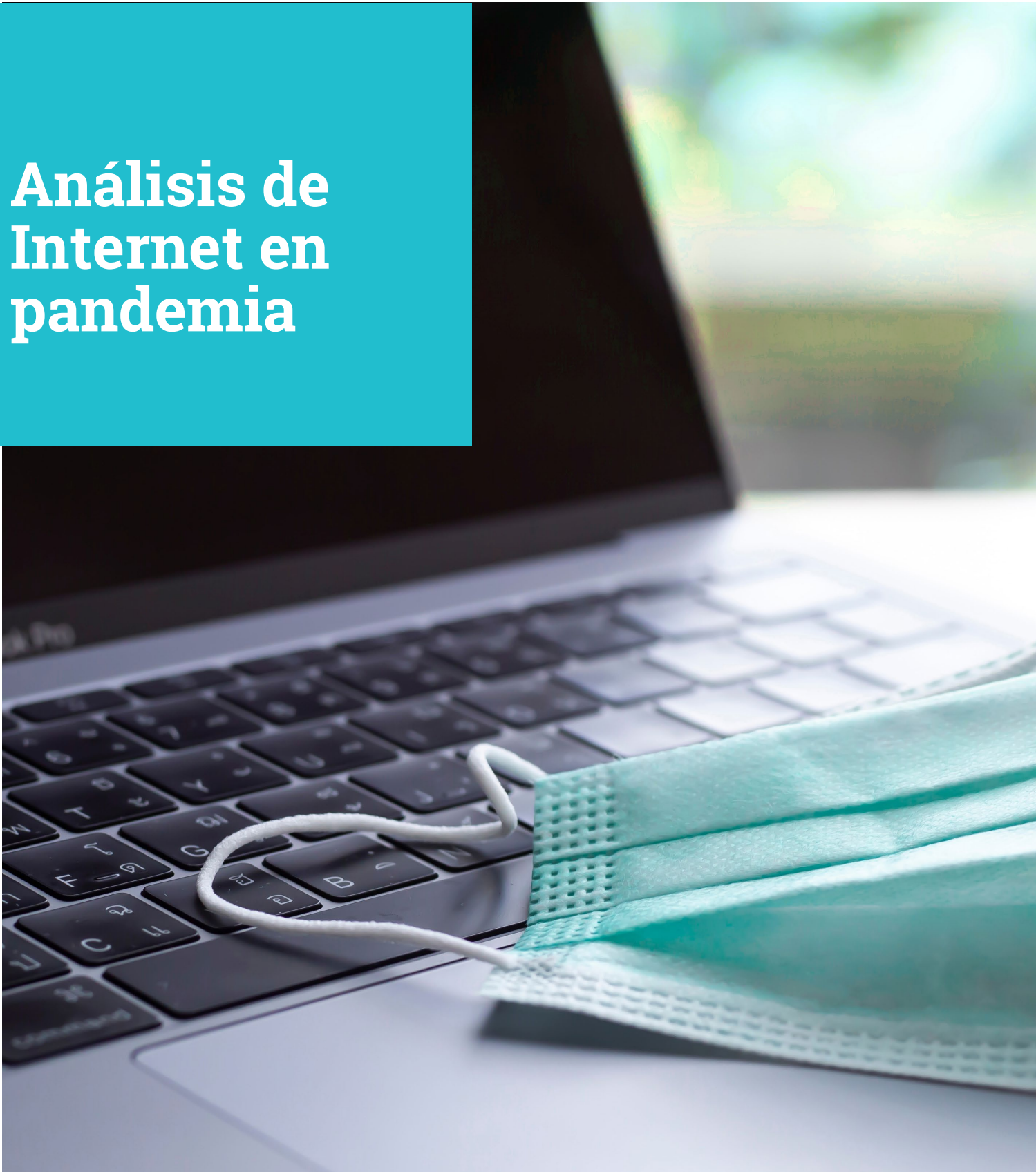




Análisis de Internet en pandemia





Equipo PePa Ping - NIC Labs



JAVIERA BERMÚDEZ
Estudiante de pregrado



JAVIER BUSTOS
Director NIC Labs



MAITE GONZÁLEZ
Ingeniera de investigación



DIEGO MADARIAGA
Estudiante de Doctorado



JAVIER MADARIAGA
Estudiante de pregrado



GABRIELA MENDOZA
Ingeniera de investigación



ALONSO REYES
Estudiante de Magíster



JUAN MANUEL SÁEZ
Estudiante de pregrado



LUCAS TORREALBA
Estudiante de pregrado



MARÍA JOSÉ VILCHES
Diseñadora multimedia



El presente estudio toma los datos recolectados por la aplicación PePa Ping¹, construida para medir la calidad de servicio de Internet y que toma periódicamente datos de uso de aplicaciones, y cuyo análisis nos sirvió para medir la diferencia en la movilidad de los estudiantes de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile durante la pandemia de COVID-19 el primer semestre de 2020.

Monitoreo de redes

“Lo que no se define no se puede medir. Lo que no se mide, no se puede mejorar. Lo que no se mejora, se degrada siempre”, declaró a fines del siglo XIX Lord William Thomson Kelvin, y su cita sigue aún siendo válida en nuestros días, especialmente en lo que a calidad de servicio de Internet se refiere. ¿Cómo se puede mejorar la calidad de Internet si no se mide?

En Chile se han realizado bastantes proyectos de mediciones de calidad como Adkintun [1], su versión Adkintun Mobile [2, 3] y Yafün [4], varios de ellos tratando de responder las mismas preguntas: *“Si queremos mejorar la calidad del Internet chileno ¿qué medir? y ¿cómo medirlo?”*. Es por esta razón que la medición constante (llamado *monitoreo*) de Internet ha sido un problema de estudio permanente desde inicios del siglo XXI.

El monitoreo de redes como Internet se puede dividir en dos grandes conjuntos: pasivo y activo. Según el RFC 7799 [5], los métodos pasivos tienen las siguientes características:

- i) están basadas únicamente en observaciones de un flujo de paquetes de interés no perturbado ni modificado,
- ii) depende de la existencia de uno o más flujos de paquetes para suministrar la información de interés, y
- iii) depende de la presencia del flujo de paquetes de interés en uno o más de los puntos de observación designados.

Por otro lado, los métodos activos (como los usados en *speedtests*) inyectan tráfico a la red hacia distintos puntos de interés y con eso realizan sus mediciones, algo no practicable en conexiones móviles con cuotas de MB de descarga mensuales.

Los sistemas operativos de dispositivos móviles no permiten, por seguridad, muchas cosas que los de computadores de escritorio sí permiten. Entre ellas, por ejemplo, está la habilidad de chequear el tráfico de salida y entrada al dispositivo. Para ello se utiliza entonces un esquema alternativo llamado “Auto-VPN”.

En una conexión normal a Internet, usuarios y empresas navegan la red para acceder a distintos servicios, como buscadores, redes sociales, o servicios propios de la empresa (por ejemplo, de finanzas). El usuario desde su casa no puede acceder a los servicios que son

propios de la empresa por razones de seguridad (ver Figura 1).

Una conexión VPN (por su nombre en inglés Virtual Private Network) es un túnel seguro usado para extender una red local (generalmente de una Empresa) hasta donde estemos ubicados. Toda la comunicación viaja encriptada por Internet, y para efectos prácticos uno estaría ubicado dentro de la red local empresarial y todo el tráfico de red se manejaría como tal (ver Figura 2).

Una conexión “Auto-VPN” entonces le hace creer al dispositivo que está enviando información a un servidor lejano, cuando en realidad se lo envía al mismo dispositivo antes que los datos vayan a Internet. Como el tráfico viaja (generalmente) encriptado lo único que uno puede “ver” son los encabezados de paquetes (usados para control y ruteo) y no el contenido, así que la verifica-

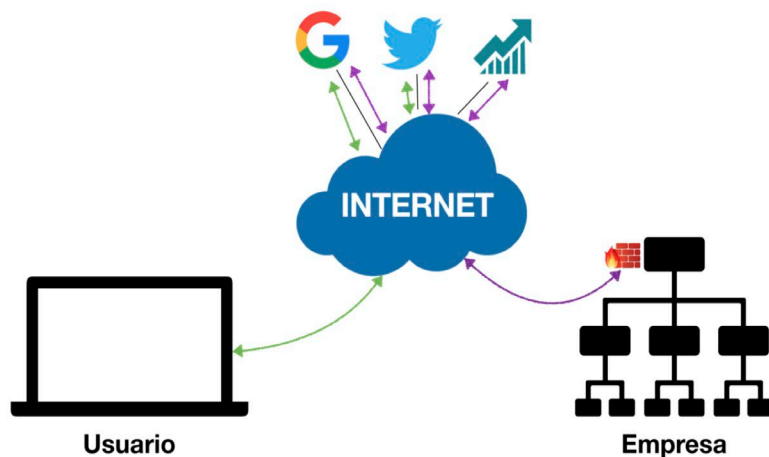


Figura 1. Conexión a Internet normal.

1 | <http://www.niclabs.cl/pepa>

ción sería la misma que usan los proveedores de Internet (ISPs) para dirigir los datos. Este esquema ha resuelto el problema de monitoreo pasivo en dispositivos móviles, y ha sido utilizado en aplicaciones académicas de monitoreo como Haystack [6], PrivacyGuard [7] y PePa Ping [8]. En la Figura 3 se presenta el esquema que utiliza PePa Ping.

Periodic Passive (PePa) Ping

Una de las principales preocupaciones a la hora de desarrollar un sistema de monitoreo de Internet, especialmente si se usará en dispositivos móviles, es la sobrecarga que puede producir en CPU y consumo de batería. Ésta es la razón de elegir un sistema de monitoreo periódico en lugar de uno basado en eventos. Con suficiente periodicidad podemos obtener una buena representación del tráfico de la red sin sobrecargar el dispositivo.

Usando un esquema de "Auto-VPN" desarrollamos desde el año 2019 un sistema de monitoreo Periódico y Pasivo llamado *PePa Ping*, un sistema capaz de tomar medidas periódicas similares a Ping (RTT, *jitter* y número de paquetes perdidos) a través de supervisión pasiva de conexiones activas y sin inyectar ningún nuevo paquete en la red, contrario al estándar Ping y sus paquetes ICMP.

PePa Ping se ejecuta de forma periódica cada 15 minutos, para monitorear el tráfico de Internet durante 1 minuto. A lo largo de cada ejecución, PePa Ping recopila información de todas las conexiones TCP y UDP del dispositivo. Además, recopila información contextual que permite una mejor comprensión de las condiciones en las que las conexiones generaron el tráfico móvil.

La lista muestra lo que se recopila como información contextual:

“Lo que no se define no se puede medir. Lo que no se mide, no se puede mejorar. Lo que no se mejora, se degrada siempre”, Lord William Thomson Kelvin.

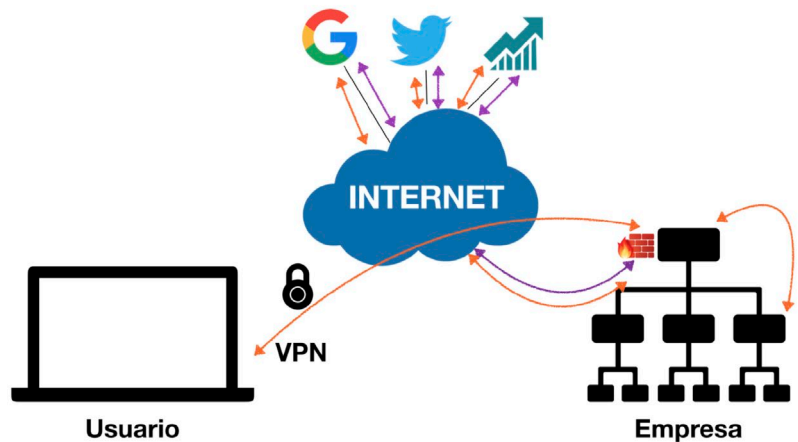


Figura 2. Conexión a Internet mediante VPN de empresa.

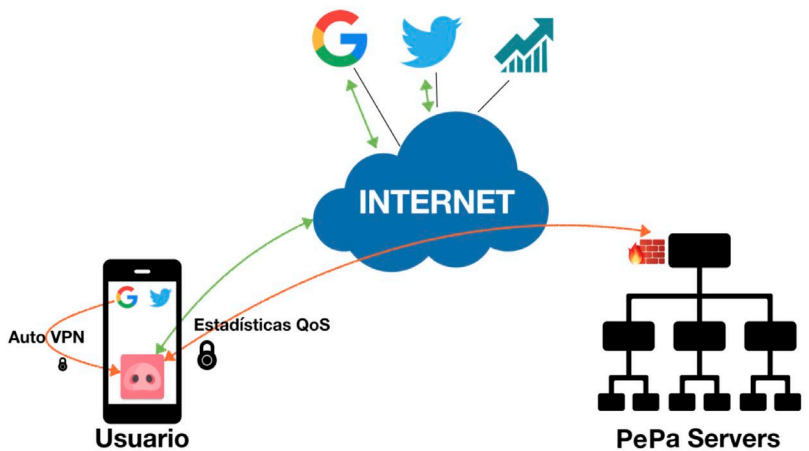


Figura 3. Conexión a Internet mediante Auto-VPN de PePa Ping.

- Tipo de conexión (WiFi o red móvil).
- Tecnología de red (UMTS, LTE, etc.).
- Indicador de intensidad de señal recibida (RSSI).
- Localización de usuario: Altitud, Longitud, Latitud, Precisión.
- Información de antena: Cell ID (CID), Location Area Code (LAC), Mobile Network Code (MNC), Mobile Country Code (MCC).
- Sim ISP (proveedor de telefonía/internet).
- Tiempo actual con precisión de segundos.

Para mejorar la experiencia de uso, PePa provee además una grata interfaz de usuario, como se muestra en la Figura 4.



Análisis de los datos

La pandemia del año 2020 nos sorprendió con el período de pruebas en usuarios reales de PePa Ping, en la cual distribuimos la aplicación principalmente entre estudiantes de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile, para probar su desempeño *in situ*. Es por esta principal razón que, aunque nuestros datos pueden no ser representativos para todo el país, si lo pueden ser para el conjunto de estudiantes de la Facultad.

Es importante mencionar que, con el fin de mantener las **consideraciones éticas** de las mediciones de la red y la gestión de los datos de los usuarios, este estudio se llevó a cabo asegurando que los usuarios de la aplicación PePa Ping estén en pleno conocimiento de cómo se utilizarán sus datos y quién accederá a ellos y los utilizará. Para ello so-

licitamos a los participantes que nos otorguen permisos para utilizar la información recopilada para actividades de investigación relacionadas con el proyecto antes de instalar la aplicación. Las mediciones recopiladas están debidamente anonimizadas y sólo los investigadores afiliados a este proyecto pueden acceder a los datos.

A partir de los datos recolectados podemos hacer entonces un estudio de uso y movilidad de dispositivos móviles de nuestros estudiantes y notar el cambio que el COVID-19 generó en sus rutinas. Dado que los datos no están distribuidos de forma normal, no podemos usar el promedio de ellos como un buen estimador. Además, lo que nos interesa es ver tendencias temporales de los datos, por lo tanto calcularemos el total de *Kbytes* por usuario y por semana. A partir de ahí, estudiaremos cómo se desplazó **la mediana** de esos datos a lo largo del año.

Demografía

Los primeros gráficos muestran que demográficamente nuestros datos corresponden con nuestros usuarios: estudiantes entre 20 y 25 años. Se evidencia en el uso de sus redes sociales favoritas (entre las cuales claramente **no está** Facebook ni Twitter, ver Figura 5) y el uso de su aplicación de *streaming* favorita (Twitch, donde se ven, entre otras cosas, eventos de juegos en línea, ver Figura 6). Notar que en los gráficos se acumulan los *Kbytes* semanales para mostrar además la proporción de uso de las distintas aplicaciones.

Como usuarios de aplicaciones de pedidos (*delivery*), nuestros jóvenes estudiantes bajan su uso durante la época de clases, salvo en el período de exámenes en el que probablemente sus necesidades alimenticias crezcan mientras el tiempo para cocinar decrece (ver Figura 7).

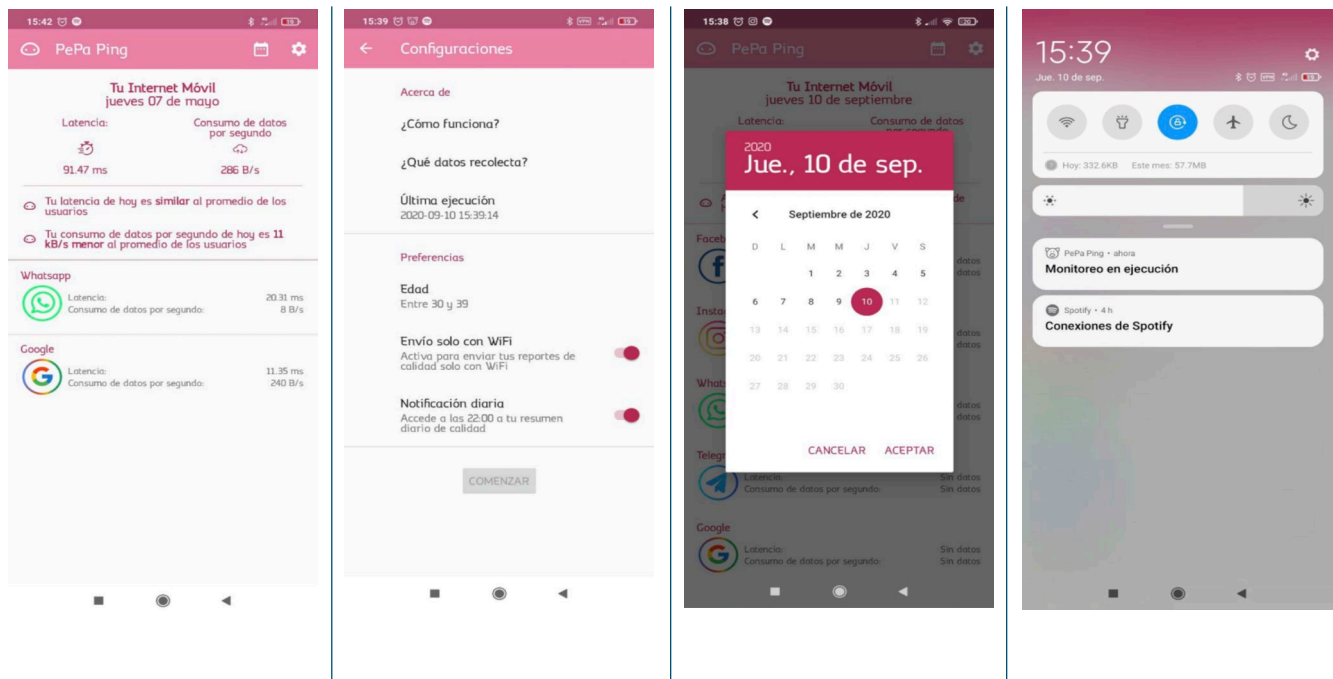


Figura 4. Interfaz de Pepa Ping.

Vida de estudiante

La Universidad de Chile utiliza como Sistema de Información Docente la plataforma U-Cursos. En ella los estudiantes encuentran sus horarios, calendarios de eventos, evaluaciones y foros de comunicación con compañeros y cuerpo docente. Su uso reflejado en la Figura 8 está dentro de lo esperado: muy bajo durante vacaciones, creciente hasta la primera evaluación, luego una baja, y vuelta a crecer para la segunda evaluación y mucho para examen y fin de semestre.

Debido a la pandemia durante el 2020 no se realizaron clases presenciales y la Universidad dejó libertad a los distintos departamentos para salvar la continuidad de enseñanza. Mientras algunos docentes lograron rediseñar sus cursos para educación online (cápsulas de video de cinco minutos, trabajo personal, lecturas, acompañamiento via email o foros de U-Cursos y sesiones de consultas) utilizando entre otras la plataforma que la Universidad dispone para ello (EOL, <http://eol.uchile.cl>), otros optaron para realizar sus clases cuasinormalmente: esta vez frente a una cámara realizaban la docencia como si estuviesen en una sala de clases: clases expositivas de 1 a 1½ horas transmitidas por Zoom o Google Meet y consultas de los estudiantes recibidas por el mismo medio en vivo y en directo. La interacción en este segundo caso se ve claramente reflejada en la Figura 9, donde se presentan las semanas en que hubo clases y las semanas en que hubo descanso.

Movilidad

PePa Ping captura datos ambientales cada vez que se activa, por lo tanto a partir de esa información se puede saber qué porcentaje de los usuarios estuvo en el mismo lugar durante el día, o en dos o más lugares. Note que si la persona se encontraba en el mismo lugar entre dos activaciones de PePa Ping, se considerará que no hubo desplazamiento.

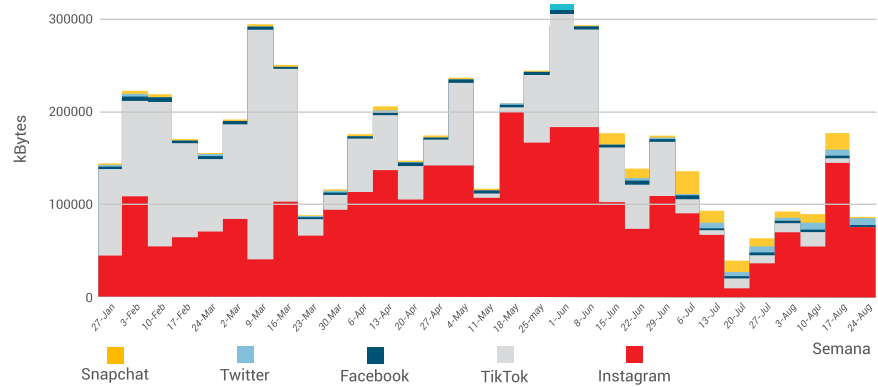


Figura 5. Mediana de uso de aplicaciones de redes sociales.

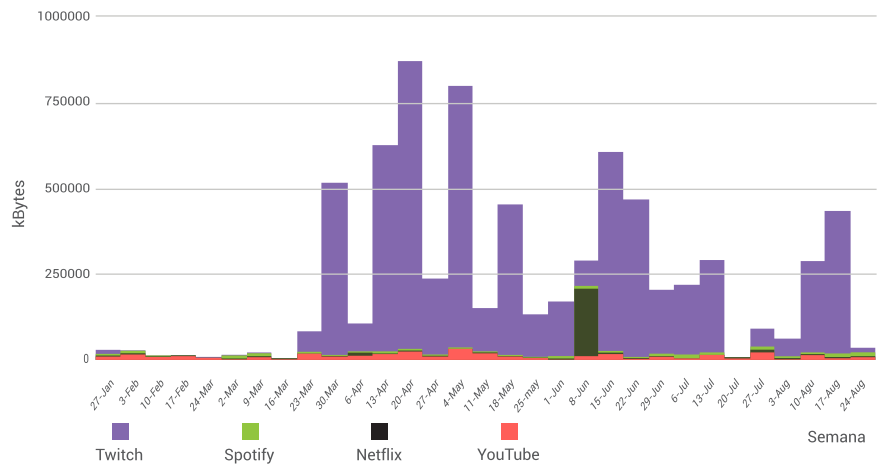


Figura 6. Mediana de uso de aplicaciones de streaming.

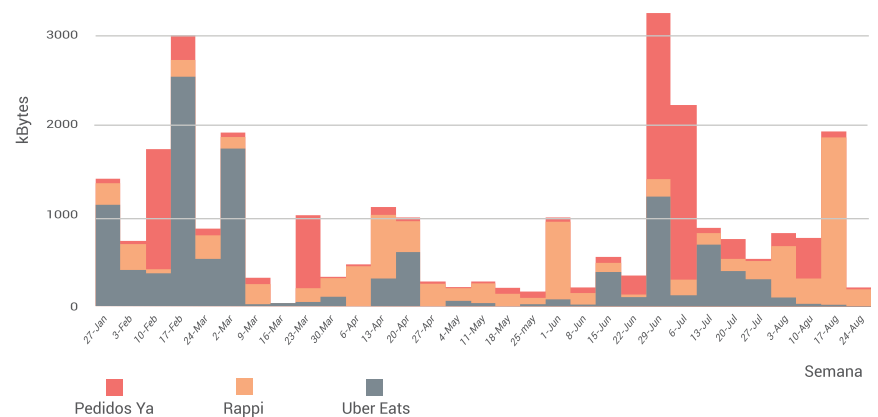


Figura 7. Mediana de uso de aplicaciones de delivery.



El uso [de U-Cursos] está dentro de lo esperado: muy bajo durante vacaciones, creciente hasta la primera evaluación, luego una baja, y vuelta a crecer para la segunda evaluación y mucho para examen y fin de semestre.

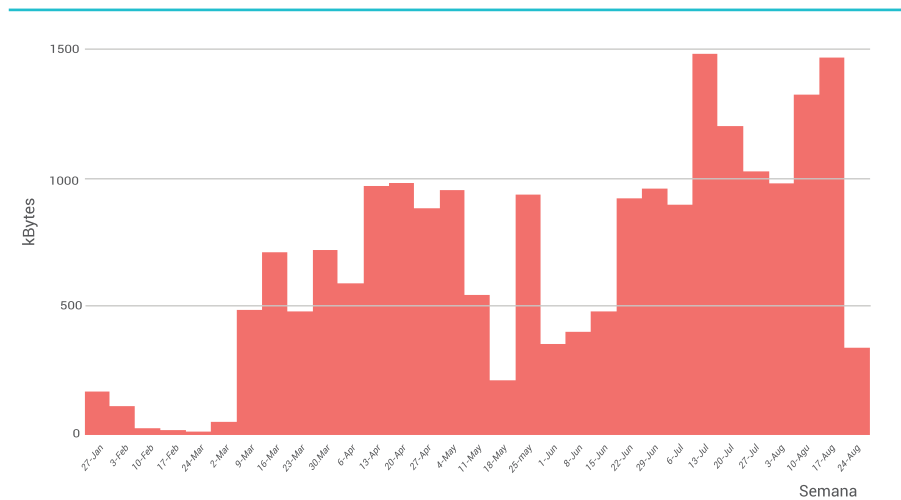


Figura 8. Mediana de uso de U-Cursos.

to aún cuando el usuario haya ido a otro lugar en el intertanto.

El análisis de movilidad se realiza, entonces, utilizando algoritmos de *clustering* de ubicación, y si los lugares están apartados por más de 200 metros se consideran como lugares distintos.

Los resultados de ejecutar esos algoritmos sobre nuestros datos se muestran en la Figura 10, donde se puede apreciar que al momento de empezar los confinamientos en Santiago debido a la pandemia por COVID-19 en Chile más del 75% de los usuarios redujeron su movilidad a **un sólo lugar**. Esa información se puede complementar con el análisis de uso de aplicaciones de transporte, especialmente TransApp que entrega información del transporte público de Santiago, y las cuales decaen intensamente después de la segunda semana de marzo (ver Figura 11).

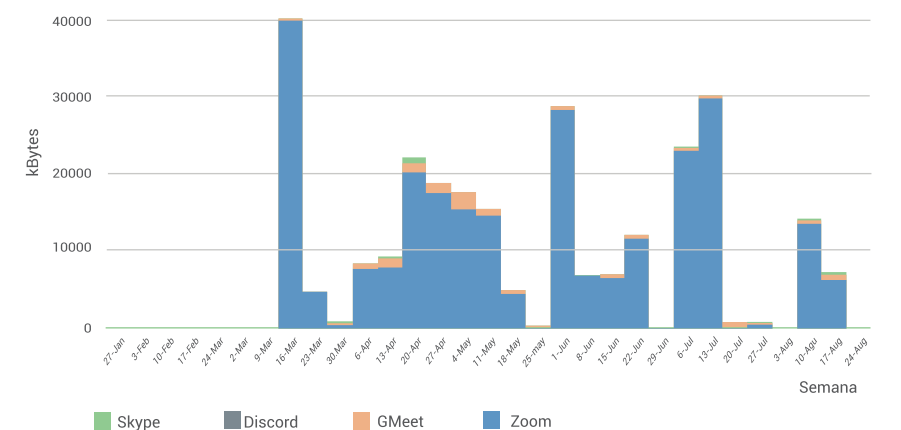


Figura 9. Mediana uso de aplicaciones de videoconferencia.

Conclusiones

El presente estudio toma los datos recolectados por la aplicación PePa Ping, analizando el uso de Internet por estudiantes de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile y cuyo análisis nos sirvió para

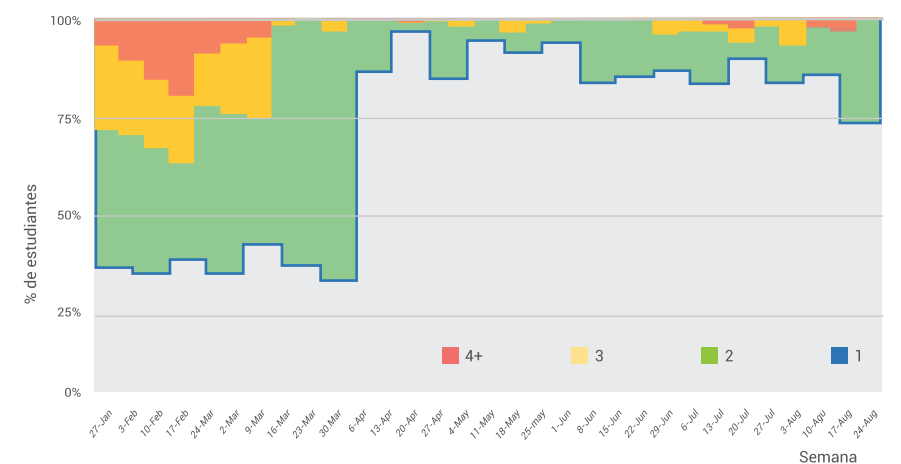


Figura 10. Porcentaje de la población ubicada en 1, 2, 3 y 4+ lugares.

Al momento de empezar los confinamientos en Santiago [...] más del 75% de los usuarios redujeron su movilidad a un sólo lugar.

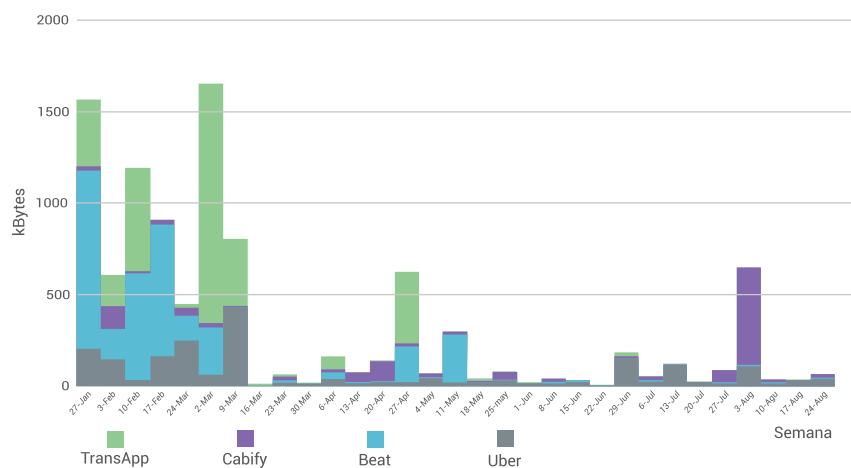


Figura 11. Mediana uso de aplicaciones de transporte.

medir el impacto de la pandemia de COVID-19 en la movilidad de estudiantes, durante el primer semestre de 2020.

Los datos muestran claramente que la movilidad se redujo en más del 75% de los casos.

Para apoyar el proyecto PePa Ping considere instalar la aplicación actualmente disponible para dispositivos Android (próximamente en iOS), para aportar a la investigación de la calidad de Internet móvil en el país. Más información acerca de PePa Ping puede encontrar en el enlace <http://www.niclabs.cl/pepa>. ■

REFERENCIAS

- [1] "Adkintun: SLA monitoring of ISP broadband offerings". J. Bustos-Jiménez, V. Ramiro, F. Lalanne, y T. Barros. *Proceedings of 27th International Conference on Advanced Information Networking and Applications Workshops*. IEEE. Páginas 1445–1449. 2013.
- [2] "Adkintun Mobile: Towards using personal and device context in assessing mobile QoS". F. Lalanne, N. Aguilera, A. Graves, y J. Bustos. *Proceedings of International Conference on Wireless Communications and Mobile Computing*. IEEE. Páginas 49–54. 2015.
- [3] "Crowd-measuring: Assessing the quality of mobile Internet from end-terminals". G. Hourton, G. Del Canto, J. Bustos-Jiménez y F. Lalanne. *Proceedings of 6th International Conference on Network Games, Control and Optimization*. IEEE. Páginas 145–148. 2012.
- [4] "Analyzing the Robustness of the Chilean Optical Network". I. Bachmann, J. Bustos-Jiménez, J. Piquer, C. Saint-Pierre, M. Rivas y M. Valenzuela. *Proceedings of ACM SIGCOMM 2nd Workshop on Optical Systems Design (OptSys)*. 2020.
- [5] "Active and Passive Metrics and Methods (with Hybrid Types In-Between)". A. Morton. IETF Informational RFC 7799. 2016
- [6] "Haystack: In Situ Mobile Traffic Analysis in User Space". A. Razaghpahan, N. Vallina-Rodríguez, S. Sundaresan, C. Kreibich, P. Gill, M. Allman y V. Paxson. arXiv. 2015.
- [7] "Privacyguard: A vpn-based platform to detect information leakage on android devices". Y. Song y U. Hengartner. *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices*. Páginas 15–26. 2015.
- [8] "PePa Ping: Android tool to take and predict periodic passive ping measurements", D. Madariaga y G. Mendoza. *Proceedings of School of Systems and Networks (SSN)*. 2018.



Celebramos 20 ediciones de Revista Bits de Ciencia, medio que nació como un espacio para estimular el interés y la discusión en torno a la ciencia de la computación.

Si te perdiste alguna edición o quieres volver a leerla, te invitamos a visitar:
www.dcc.uchile.cl/bits-de-ciencia





REVISTA DEL DEPARTAMENTO DE CIENCIAS DE LA
COMPUTACIÓN DE LA UNIVERSIDAD DE CHILE

Bits

DE CIENCIA