Random UChile

Aleatoriedad transparente en procesos públicos

CAMILO GÓMEZ

Magíster en Ciencias mención Computación por la Universidad de Chile. Coordinador Operativo de Random UChile.

cjgomez@uchile.cl

ALEJANDRO HEVIA

Doctor en Computación por la University of California, San Diego. Profesor Asociado del Departamento de Ciencias de la Computación de la Universidad de Chile. Director del Laboratorio de Criptografía Aplicada y Ciberseguridad (CLCERT). Sus áreas de investigación incluyen criptografía aplicada y seguridad computacional.

ahevia@dcc.uchile.cl

BRYAN ORTIZ

Estudiante de Magíster en Ciencias mención Computación en la Universidad de Chile. Desarrollador en el proyecto Random UChile.

bortiz@dcc.uchile.cl

RESUMEN. El "azar" (o aleatoriedad) forma parte de nuestra vida cotidiana. Un evento aleatorio puede surgir, por ejemplo, cuando participamos en un concurso o se nos asigna una tarea. Por lo general, estos eventos involucran el lanzamiento de una moneda o un dado o, en casos más complejos, una consulta a un algoritmo computacional que genera bits "aleatorios". Sin embargo, dicha consulta es opaca para el exterior y puede ser insuficiente en contextos donde se requiere mayor transparencia. Los procesos públicos, en particular, como la asignación de auditorías o la selección de vocales de mesa, hoy en día parecen requerir plena confianza en la entidad que realiza el sorteo, a fin de garantizar que el resultado no haya sido manipulado. Un faro de aleatoriedad es un servicio que busca cambiar este escenario, aportando con aleatoriedad pública y verificable. Un faro produce de forma periódica valores impredecibles pero públicamente disponibles y verificables. Desde el 2017, el Laboratorio de Criptografía Aplicada y Ciberseguridad, cuenta con Random UChile, un faro de aleatoriedad desde Chile para el mundo.

¿Qué es la aleatoriedad pública y para qué sirve?

Probablemente, todos hemos participado alguna vez en un sorteo. La emoción y las ganas de participar residen en que todos los participantes tenemos las mismas posibilidades de ganar. Cuando nos involucramos en una rifa en el colegio, un lanzamiento de ruleta en un casino, un concurso en redes sociales o incluso al decidir quién paga la cuenta en un almuerzo, deseamos que el resultado sea verdaderamente al azar. Vivimos constantemente rodeados de procesos aleatorios, sistemas cuyo resultado final debiese ser impredecible. De hecho, dicha impredecibilidad intrínseca es lo que a menudo nos impulsa a participar en ellos, es la garantía que nos hace ver el proceso como justo.

Ahora bien, notemos que si el proceso aleatorio es totalmente impredecible, significa que cualquier resultado es plausible. Entonces, ¿cómo podemos asegurar que el resultado no fue elegido anteriormente de forma deliberada? ¿Es posible distinguir entre un resultado aleatorio o un resultado definido "a dedo"? En escenarios como los antes descritos, esto puede llegar a ser un problema serio. Tal vez @pedritomaster77 no ganó el sorteo de esos chocolates por mera suerte, después de todo. Esto se vuelve más crítico en procesos de alta connotación, con

implicancias más importantes, como ser elegido para que auditen mis impuestos, ser seleccionado como vocal de mesa o perder una adjudicación aleatoria de fondos públicos. Las consecuencias de la desconfianza en tales escenarios pueden conllevar tiempo y costos significativos, además de un daño permanente a la confianza y a la reputación pública. Traer transparencia a dichos procesos de selección se ve necesario, pero a priori no es claro cómo lograrlo.

La aleatoriedad pública y verificable apunta a resolver dicho problema. Permite que cualquier persona pueda verificar el resultado y quedar así convencida de la correctitud del proceso —que nadie "metió mano". El proyecto Random UChile se dedica a esta tarea, mediante el servicio denominado *faro de aleatoriedad*. El faro genera, cada minuto, un valor público e impredecible de 512 bits de largo, denominado *pulso aleatorio*.

Si el proceso aleatorio es totalmente impredecible, significa que cualquier resultado es plausible. Entonces, ¿cómo podemos asegurar que el resultado no fue elegido anteriormente de forma deliberada?

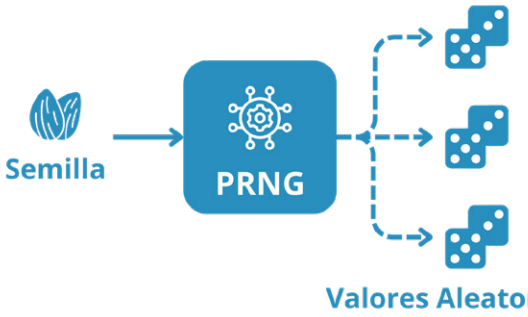

Figura 1. Descripción gráfica del funcionamiento de un generador pseudoaleatorio PRNG *(pseudorandom number generator)*.

¿Cómo funciona? Semillas y generadores pseudoaleatorios

Cuando uno solicita a un computador un número aleatorio, en realidad estamos solicitando un valor *pseudoaleatorio*. Este valor *pseudoaleatorio* es generado utilizando un algoritmo de generación pseudoaleatoria (o *Pseudo-Random Number Generator, PRNG*), el cual utiliza un valor inicial denominado semilla (ver Figura 1). Si el valor específico de la semilla es secreto, los números generados por dicho generador serán *impredecibles*. Sin embargo, si repetimos el mismo proceso en otro equipo, utilizando la misma semilla, los resultados generados serán exactamente los mismos. Por lo tanto, si se conoce la semilla utilizada, la impredecibilidad es sólo una ilusión. En particular, podría ocurrir que quien realiza el sorteo eligiese una semilla conveniente (previamente probada) que produzca el resultado que desea obtener. La elección y uso de una buena semilla es entonces crucial. Para usar un PRNG, los computadores intentan generar una buena semilla recolectando información aleatoria de su ambiente como, por ejemplo, los milisegundos entre teclas presionadas, la velocidad del ventilador, o la frecuencia de llegada de los paquetes de red, entre otros.

La idea de la aleatoriedad verificable es simple: en vez de establecer su propia semilla, la persona a cargo del sorteo se compromete a utilizar como semilla un futuro pulso aleatorio del faro de aleatoriedad. Si el valor de dicha semilla es impredecible, el resultado será también impredecible. Más aún, si el valor de la semilla del faro es subsecuentemente publicado, cualquiera podrá verificar dicho resultado. Esto impide que, de antemano, la persona a cargo del sorteo ajuste el resultado a un valor específico.

Origen de Random UChile

El proyecto Random UChile surge el año 2017, buscando implementar un servicio del tipo *randomness beacon*, siguiendo los lineamientos del National Institute of Standards and Technology (NIST), la principal agencia de estándares y tecnología de los Estados Unidos. El equipo investigador de NIST ya había puesto en marcha el primer faro de aleatoriedad un poco tiempo antes y buscaba levantar otros faros de aleatoriedad, en otras partes del mundo, con el objetivo de formar una red mundial distribuida y resiliente. Acogiendo esta llamada, el Laboratorio de Ciberseguridad y Criptografía Aplicada de la Universidad de Chile (CLCERT), en cooperación con NIST, crea Random UChile, el cual se diferencia del proyecto de NIST en dos aspectos importantes: (1) prioriza el uso de fuentes externas de

Figura 2. Un ejemplo de TRNG: un generador cuántico de números aleatorios (QRNG) con interfaz PCIe.

Figura 3. Recolección y procesamiento de entropía a partir de fuentes externas.

aleatoriedad, y (2) incorpora en su desarrollo el identificar activamente escenarios donde utilizar la aleatoriedad pública generada.

Fuentes de aleatoriedad

El faro de aleatoriedad basa su funcionamiento en una recolección continua de datos aleatorios a partir de distintas fuentes de entropía. La *entropía* es una medida matemática para saber "qué tan aleatorio es un proceso", y se mide en bits. Por ejemplo, el lanzamiento de una moneda sin sesgo es un proceso aleatorio que contiene 1 bit de entropía (cara o sello). Por ello diremos que una persona que lanza una moneda sin sesgo, es una fuente que genera 1 bit de entropía cada vez que un lanzamiento ocurre. En general, las fuentes a utilizar son clasificadas en internas y externas.

Las fuentes internas son aquellas que residen dentro del servicio, como el dado por un módulo de *hardware* denominado *true random number generator* (TRNG), presentado en la Figura 2. Usando un proceso cuántico basado en la detección de fotones [1], este dispositivo produce un valor aleatorio cada vez que es invocado. Otra fuente interna es un *hardware security module* (HSM), el cual contiene un generador basado en procesos físicos para generar valores aleatorios, en forma similar.

Las fuentes externas residen fuera y son completamente independientes del servicio. Por ejemplo, el faro recolecta entropía a partir del Centro Sismológico Nacional, vía su sitio web, observando las características del último sismo de magnitud superior a 2.5 registrado en Chile. También, el faro obtiene aleatoriedad del *streaming* de la Radio Universidad de Chile, capturando extractos de aproximadamente 5 segundos (aprox. 100 *kilobytes*) en base a un criterio de recolección fijo. La última fuente externa de aleatoriedad del faro es la *blockchain* de la criptomoneda Ethereum, desde donde extrae el *hash* asociado al último bloque registrado en la cadena. Para cada una de las fuentes, se calcula el valor procesado como el hash del valor "bruto" obtenido desde la fuente. La Figura 3 ilustra el proceso de extracción de entropía desde las tres fuentes externas.

El uso de fuentes externas entrega impredictibilidad. En palabras simples, "si no se puede predecir lo que se extrae de las fuentes, entonces no se puede predecir lo generado por el faro de aleatoriedad". La información de sismos es (bajo ciertas condiciones) considerada impredecible, y similarmente la del último bloque de una *blockchain* como la de Ethereum. El uso de la radio como fuente externa viene de la dificultad de predecir exactamente los bytes que emitirá el streaming durante una transmisión en vivo. Formalizar y cuantificar esta "dificultad de predecir" es un desafío importante y toma la forma de una estimación de la entropía asociada a cada fuente, por minuto. Este indicador, actualmente un trabajo en desarrollo, busca robustecer el sistema dando garantías en tiempo real del nivel de aleatoriedad de cada pulso generado.

Los datos aleatorios recolectados, tanto internos como externos, son combinados mediante un algoritmo criptográfico, para generar un único valor de salida, denominado *pulso aleatorio*. La impredecibilidad de todos estos eventos proporcionan una garantía de aleatoriedad (alta entropía) al pulso generado cada minuto.

Además de la generación de valores aleatorios y públicos, el faro de aleatoriedad implementa mecanismos que permiten la inmutabilidad de los valores generados, ofreciendo la posibilidad de verificar, en el futuro, la correcta generación de valores pasados, así como también poder consultar cualquier valor generado desde el principio de los tiempos.

El uso de fuentes externas entrega impredictibilidad. En palabras simples, "si no se puede predecir lo que se extrae de las fuentes, entonces no se puede predecir lo generado por el faro de aleatoriedad".

¿Dónde utilizar la aleatoriedad pública?

Aplicaciones simples

En la vida cotidiana, la aleatoriedad producida por el faro puede ser utilizada para escoger una opción al azar dentro de un grupo de opciones, realizar sorteos por medio de redes sociales, o incluso como un dado de *N* caras en juegos de mesa. La principal ventaja de optar por esta fuente de aleatoriedad es que brinda legitimidad al proceso, permitiendo que cualquier persona pueda verificar el resultado. En el sitio web de Random UChile contamos con estas y otras aplicaciones, permitiendo la verificación del resultado por medio de un enlace o código QR.

Investigaciones científicas

El uso de aleatoriedad en investigaciones científicas es una práctica común para llevar a cabo experimentos o muestreos representativos. Sin embargo, la utilización de aleatoriedad "opaca", o no transparente, dificulta la reproducción de resultados precisos. Peor aún, existe la posibilidad de que los datos utilizados hayan sido seleccionados de manera sesgada con el propósito de alterar los resultados en favor de ciertas conclusiones. Un ejemplo claro de esto se observa en la partición de un dataset para el entrenamiento y evaluación de modelos de *machine learning*, donde la falta de aleatoriedad puede influir en los resultados finales de manera injusta. En este contexto, utilizar una fuente de aleatoriedad verificable y pública en el proceso de selección implica tomar una posición *activamente transparente*, al brindar una validación amplia y pública al proceso de experimentación, lo cual ayuda a mitigar posibles sesgos o manipulaciones indebidas de los datos.

Asignación de jueces y resolución de conflictos

La aleatoriedad verificable tiene un papel crucial en la prevención activa de conflictos de intereses en procesos públicos, evitando así los gérmenes de la desconfianza en tales procedimientos. Tomemos, como ejemplo, la asignación de jueces en casos judiciales: la posibilidad de que un juez tenga conexiones personales con una de las partes puede comprometer su imparcialidad. Sin embargo, al emplear la aleatoriedad verificable en este proceso, se garantiza transparencia y se asegura a la ciudadanía que la asignación se realizó de manera imparcial y objetiva. Un efecto similar tendría la asignación de jueces en la resolución de conflictos entre distintas entidades que desean, por ejemplo, registrar el mismo dominio web. En ambos casos, la presencia de un mecanismo público y verificable de aleatoriedad fortalece la legitimidad del proceso aleatorio y, en consecuencia, la confianza en su integridad.

Trabajos desarrollados

Dentro de las principales motivaciones de la generación de aleatoriedad verificable está el garantizar la correctitud de procesos públicos, particularmente aquellos que utilizan el azar para la toma de decisiones de alta connotación. Random UChile ha sido partícipe de tres proyectos de esta categoría. El primero, está relacionado con la selección de auditorías fiscales en la principal entidad fiscalizadora estatal chilena. El segundo, buscó modificar la selección de vocales de mesa en votaciones, y, el tercero, dió credibilidad a la selección de ciudadanos para un proceso de deliberación pública. Los tres proyectos se describen a continuación.

Selección de Auditorías (CGR). La Contraloría General de la República (CGR) es el organismo encargado de fiscalizar el gasto fiscal en el país. Como tal, debe seleccionar periódicamente, y aplicando un factor aleatorio, a los empleados públicos que deben rendir auditoría y ser fiscalizados. Previamente a nuestro proyecto, al utilizar un método cerrado, no existían garantías de la aleatoriedad del proceso, por lo que los resultados bien podrían ser desacreditados como persecución política. En el marco del trabajo de tesis de Constanza Csori [2], Random UChile implementó un prototipo que apoyaba el proceso llevado a cabo por la CGR, seleccionando empleados públicos para ser auditados con base a la aleatoriedad pública del faro.

Vocales de Mesa (SERVEL). En procesos electorales recientes, la selección de vocales de mesa no ha estado exenta de polémica [3]. Si bien se supone que son seleccionados incorporando el azar dentro de cada junta electoral, en muchas de ellas el proceso no es transparente ni público. Considerando que una selección partidista para estos cargos puede generar desconfianza en el proceso e incluso abrir la puerta a corrupción, resulta imperativo brindar transparencia al proceso. En el marco de la memoria de título de Franco Pino [4], y con la colaboración de la junta electoral de Macul, se desarrolló un prototipo de selección verificable de voca-

Figura 4. Ilustración de "Las y Los 400: Chile Delibera", por Fundación Tribu.

Figura 5. Diseño hecho por Cloudfare en 2019 para los miembros fundadores de la Liga de la Entropía. De izquierda a derecha se encuentran *ChaChaRand* de Kudelski Security, *Interplanetary Girl* de Protocol Labs, *LavaRand* de Cloudfare, *Seismic Girl* de Random UChile, y *URand* de EPFL.

les de mesa, utilizando como fuente de aleatoriedad el faro de Random UChile. La solución propuesta dotaría de transparencia al proceso de selección y ofrecería además un método de verificación público, de modo que cualquier ciudadano podría confirmar el nombramiento justo de los vocales de una mesa.

LXS 400. En 2019, la Fundación Tribu, en conjunto con el Centro para la Democracia Deliberativa de la Universidad de Stanford, impulsaron un proceso de deliberación pública donde alrededor de 400 personas discutieron temáticas de contingencia nacional (ver Figura 4). El proceso de selección se llevó a cabo utilizando el faro de aleatoriedad de Random UChile, apuntando a obtener independencia, transversalidad y representatividad de los ciudadanos elegidos. En particular, se realizaron dos sorteos: primero, se seleccionó aleatoriamente 30.000 manzanas a lo largo de todo el territorio nacional, cuyos resultados se encuentran públicamente disponibles [5]; el segundo sorteo se realizó en conjunto con el Centro de Microdatos de la Universidad de Chile, el cuál definió las viviendas específicas que fueron seleccionadas para ser invitadas a participar de esta instancia. Este último paso se realizó de forma privada, con el fin de mantener la privacidad de las personas invitadas.[1]

Colaboraciones actuales

NIST

NIST publicó en 2019 el borrador de la referencia para faros de aleatoriedad [6], definiendo una versión 2.0 que apunta, principalmente, a la interoperabilidad entre distintos faros de esta índole. En la actualidad, tanto Random UChile como el Instituto Inmetro de Brasil y NIST, nos encontramos en un proceso de colaboración continua con el fin de impulsar una nueva versión de este documento. Esta nueva versión apunta a definir un estándar a largo plazo para guiar el surgimiento de nuevos faros de aleatoriedad en el mundo.

DRAND - Liga de la Entropía

DRAND (pronunciado *di-rand*) es un faro de aleatoriedad distribuido desarrollado en el laboratorio DEDIS de École Polytechnique Fédérale de Lausanne (EPFL). Este faro tiene la particularidad de que la generación de aleatoriedad se realiza de manera distribuida, donde cada uno de los participantes interactúa con los otros para ir generando, de manera periódica valores aleatorios. La instancia más importante de DRAND es La Liga de la Entropía, la cual está conformada por un conjunto de servidores alrededor del mundo, administrados cada uno por diversas organizaciones, que corren una instancia del protocolo DRAND. Cada uno de los nodos participantes contribuye con sus fuentes individuales de alta entropía para poder así generar números aleatorios, impredecibles y verificables. Random UChile forma parte de esta red desde sus inicios en 2019 (ver Figura 5).

1 Aún así, la semilla aleatoria utilizada para esta parte era auditable en caso de que ello hubiera sido requerido y autorizado.

Tomemos, como ejemplo, la asignación de jueces en casos judiciales [...] Al emplear la aleatoriedad verificable en este proceso, se garantiza transparencia y se asegura a la ciudadanía que la asignación se realizó de manera imparcial.

Conclusiones finales

Muchos procesos aleatorios que encontramos en nuestra vida diaria utilizan una forma de aleatoriedad que no es fácil de verificar (valores escogidos al azar en forma opaca), lo que imposibilita disipar las dudas sobre si el resultado fue realmente aleatorio o si hubo manipulación. Esta falta de transparencia puede ser especialmente problemática en situaciones públicas o donde el resultado es de gran importancia, ya que posibles conflictos de interés pueden sembrar desconfianza tanto entre los participantes como entre los observadores del proceso. Los faros de aleatoriedad, como el que ofrece Random UChile, buscan mitigar este problema proporcionando una fuente de aleatoriedad pública y verificable, llenando un vacío existente en nuestra infraestructura digital. Su existencia provee nuevos mecanismos para proveer transparencia e imparcialidad (y finalmente legitimidad) a procesos digitales previamente opacos, algo fundamental para construir confianza en nuestra sociedad digital. ▪

BIBLIOGRAFÍA:

[1] Quantis, “Redefining Randomness, True Random Number Generator”. https://www.idquantique.com/random-number-generation/products/quantis-random-number-generator/ (accedido 22 de mayo 2024).

[2] Constanza Csori, “Achieving transparency in public decision making processes via verifiable randomness”, Repositorio de tesis UChile, 2019. https://repositorio.uchile.cl/bitstream/handle/2250/175080/Achieving-transparency-in-public-decision-making.pdf.

[3] Francisca Rivas, “Consejera del Servel explica por qué hay vocales de mesa que se han repetido tantas veces”. Bío Bío Chile. 17 de diciembre 2023. https://www.biobiochile.cl/noticias/sociedad/debate/2023/12/17/consejera-del-servel-explica-por-que-hay-vocales-de-mesa-que-se-han-repetido-tantas-veces.shtml.

[4] Franco Pino, “Elección de vocales de mesa con aleatoriedad verificable”, Repositorio de tesis UChile, 2019. https://repositorio.uchile.cl/bitstream/handle/2250/173971/cf-pino_fc.pdf.

[5] Resultados sorteo Lxs 400. Realizado 9 de diciembre 2020. https://clcert.github.io/lxs400-sitio-resultados/.

[6] John Kelsey, Luís Brandão, René Peralta, Harold Booth, “A Reference for Randomness Beacons”, Draft NISTIR 8213, 2019. https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8213-draft.pdf.