martedì 27 luglio 2010

di *Marco Calamari*

A+ A-

Commenti (10)

Cassandra Crossing/ HOT Bits of PETS

di Marco Calamari - Da Tor alla reidentificazione dei dati medici anonimizzati, passando per la geolocalizzazione dei servizi di tariffazione autostradale. La seconda puntata di un reportage atipico da PETS 2010

Roma – Incredibilmente ben 4 persone hanno manifestato qualche forma di interesse per questa cronaca "leggera" di PETS 2010 e poiché almeno due di queste non mi risultano pagate per farlo, ecco qui l'ipotizzata seconda puntata, più tecnica e non solo di "colore".

Innanzitutto vi ho detto che tra i 109 partecipanti ce ne erano almeno una dozzina del Progetto Tor? Ah no? C'erano. Se qualche potenza straniera o gli alieni avessero voluto azzerare il Progetto questa sarebbe stata un'occasione perfetta.
A giudicare da questo folto plotone ed anche dal numero di interventi in materia, Tor gode di ottima salute, il suo sviluppo procede con continuità ed alcune nuove idee "rivoluzionarie" come la tipologia di directory server e la topologia della rete sono in fase di discussione.

La maggior parte dei dibattito su Tor è attualmente incentrata sull'effettivo uso della tipologia di nodi "bridge", appositamente concepiti per contrastare l'attività di governi come quello cinese che blocca gli indirizzi dei nodi Tor pubblici per impedirne l'uso da parte dei suoi cittadini.
L'utilizzo effettivo dei nodi di tipo relay, che consumano pochissima banda perché non partecipano al traffico smistato dai normali router ma agiscono solo come punto di ingresso nascosto alla rete Tor, è stato oggetto di studi accurati che hanno permesso di scoprire anche alcune reazioni messe in atto dai gestori del "Grande Firewall cinese". "Watch the Watchmen", insomma. Se volete dare una mano direttamente a chi vive in regimi meno liberali dei nostri,

abilitate questa semplice funzionalità nel vostro nodo.

Un altro tema di discussione su Tor è stato il ridurre a due, da tre che sono adesso, i router Tor che vengono normalmente usati per realizzare una connessione. Il secondo router di una connessione, quello middleman, aumenta la sicurezza del sistema "separando" il router di ingresso da quello di uscita, ma di converso rallenta evidentemente il funzionamento e peggiora quindi la user experience; inoltre eliminare un hop permetterebbe di risparmiare il 33% di banda della rete Tor.
Tor nasce come sistema a bassa latenza per l'anonimato, per navigare insomma, perciò non può permettersi di essere troppo lento. L'eliminazione di un hop fa sperimentalmente diminuire di un buon 30% la latenza, mentre un'analisi qualitativa degli attacchi possibili contro questa modifica non rivela diminuzioni marcate di sicurezza. Staremo a vedere.

Ma non c'era ovviamente solo Tor tra i temi più interessanti. Non vi ho ancora raccontato che il PETS, come altre manifestazioni di questo tipo, è diviso in tre momenti: il Symposium che è la parte principale in cui vengono presentati i lavori più formali e soggetti a peer review; una Rump session, in cui chiunque può chiedere fino all'ultimo momento 5 minuti per parlare anche a braccio di qualsiasi cosa; la HotPETS in cui vengono presentati lavori importanti ma molto innovativi o ancora incompleti.

E proprio in quest'ultimo spazio, a cui è stata dedicata un'intera giornata e i cui atti sono pubblici e gratuiti, si sono a parer mio sentire le cose piu interessanti. Gli atti di HotPETS sono scaricabili qui e ve ne consiglio senz'altro la lettura

Un tema molto "caldo" è stato quello della privacy dei dati di geolocalizzazione (un chiodo fisso di Cassandra, come ben sapete) generati da sistemi di tariffazione del traffico autostradale. Molti paesi, tra cui l'UE, introdurranno l'obbligo di tariffare l'uso delle strade a pagamento su basi di effettivo consumo e di uso "virtuoso" (per esempio a bassa velocità o nelle ore di basso traffico). Questo richiederà l'installazione di client tipo Telepass che rilevino la posizione in tutti i veicoli, e di un sistema informatico che conoscendo le posizioni

successive del veicolo e i dati del proprietario calcoli il costo del tragitto.

Un siffatto sistema, che sarà obbligatorio per legge, dovrebbe essere ovviamente progettato "Privacy by Design", cioè decidendo l'architettura del sistema in modo da ottenere la funzionalità con la minima esposizione possibile di dati dell'utente. Ad esempio, un sistema potrebbe essere realizzato con Telepass dotati di GPS che rilevassero la posizione del veicolo e la trasmettessero semplicemente ad un sistema centrale. Quest'ultimo, conoscendo la posizione di tutti i veicoli, sarebbe in grado di calcolare i costi ma a prezzo di una potenziale esposizione (per errore, per un bug o per frode) dei dati dettagliati sulla posizione di tutti gli utenti ogni volta che hanno usato la loro auto.

Realizzare un sistema più sicuro che implementi la "Privacy by Design" implicherebbe l'uso di un'architettura diversa e non necessariamente più costosa, in cui il Telepass a bordo dell'auto sarebbe più "intelligente": il client all'inizio del viaggio riceverebbe via radio dalla rete l'elenco delle tariffe e gli altri dati amministrativi necessari, calcolerebbe lui stesso il costo del percorso e trasmetterebbe solo questo al sistema centrale. Client disonesti che volessero barare sarebbero poi rilevati con controlli a campione utilizzando ad esempio un sistema indipendente di lettura automatica delle targhe. In questo modo si renderebbe impossibile la diffusione di dati sulla geolocalizzazione dei veicoli. Ci sono dei bei soldi in ballo, e sarebbe questo il momento di prendere decisioni importanti ed assennate, evitando magari domani di dover correggere sistemi esistenti aggiungendoci funzionalità, che risulterebbero inevitabilmente meno efficaci e più costose. Privacy by Design è anche il nome di una organizzazione dedicata al tema (i canadesi sono molto avanti su questi temi, è un paese civile).

Una fatto interessante, che dovrebbe far pensare chi non si è curato di farlo prima, è l'utilizzo di nodi Tor di uscita per l'"harvesting" di dati: si tratta insomma di utilizzare il traffico in ingresso ed uscita dal nodo per effettuare delle analisi oppure addirittura per degli esperimenti. Ben due dei lavori presentati avevano utilizzato questa tecnica di raccolta dei dati, ed uno di essi non aveva nemmeno nulla a che fare con Tor o l'anonimato in Rete. Insomma, solo una strada comoda per avere dati interessanti. Niente di nuovo sotto il sole, il traffico in uscita da Tor non è cifrato, tutti lo sanno ma forse non abbastanza ne traggono sempre le debite conseguenza. *Be warned*...

Uno dei suddetti lavori era l'interessante tentativo di analizzare l'utilizzo del P2P ed in particolare di BitTorrent che alcuni utenti fanno (o tentano di fare) attraverso Tor. Le conclusioni sono che ovviamente scaricare con BitTorrent attraverso Tor è quasi sempre una cattiva idea, e che anche limitarsi alla ricerca dei tracker non offre grandi miglioramenti di privacy, in particolare se si comincia subito ad effettuare il download. Nelle conclusioni dei due interventi suddetti i relatori si sono comunque preoccupati di dare garanzie sull'uso e sulla cancellazione dei dati raccolti tramite exit node. Una specie di "*We are not evil*". Probabilmente

era anche vero, ma non mi ha convinto per niente. Di certo è stata anche una operazione borderline dal punto di vista legale.

Last but not least, un interessantissimo ed agghiacciante intervento sulla reidentificazione di dati medici e legali. Mi ha particolarmente interessato perché come forse sapete, deanonimizzazione e reidentificazione sono stati i temi guida (ancorché trattati in pochi interventi) di e-privacy 2010.
Un ottimo riassunto del settore è consultabile in questa paper di Paul Ohm. Un ricercatore olandese ha ripetuto le analisi volte ad identificare a chi si riferivano dei set di dati pubblicati per scopi di ricerca dopo essere stati anonimizzati. Ha fatto questo utilizzando però dati reali ed attuali del 16% dei cittadini olandesi, ottenuti per vie ufficiali, deanonimizzandoli ed eseguendo poi interessanti analisi statistiche sui QID (Quasi-Identifiying Data) che possono essere impiegati nei casi più comuni. Ha anche fornito un'interessante serie di dati e grafici sulla dimensione degli insiemi di anonimato. Disturbante e vivamente raccomandata lettura.

Le pubblicazioni di tutti gli interventi che ho citato sono reperibili negli atti di HotPETS. Si leggono molto bene, e se davvero questi temi vi interessano non saranno accettate scuse da coloro che non provassero almeno a scorrerli.
Per approfondire c'è poi questa esaustiva bibliografia sull'anonimato del Progetto Free Haven. Fortemente consigliata per esigenze documentative di qualsiasi livello.

Marco Calamari

Lo Slog (Static Blog) di Marco Calamari

Tutte le release di Cassandra Crossing sono disponibili a questo indirizzo

TAG: Internet, privacy, PETS 2010, anonimato, geolocalizzazione, Tor, Quasi-Identifiying Data

Risposta alla notizia

Prima di pubblicare un tuo commento assicurati che:

- sia in tema e contribuisca alla discussione in corso
- non abbia contenuto razzista o sessista
- non sia offensivo, calunnioso o diffamante

La redazione con i controlli a campione si riserva di cancellare qualsiasi contenuto ingiurioso, volgare, illegale o contrario alla policy.

Nome e cognome Fai il login o Registrati

Oggetto Emoticon e video

Testo - Anteprima - Fai l'upload di un video

caratteri disponibili: 7000

Ho letto e approvato la policy dei commenti. Il post che sto inserendo non contiene offese e volgarità, non è diffamante e non viola le leggi italiane.

Invia