Collegio Ghislieri

Convegno 2015 di Informatica giuridica

- UN SENTITO RINGRAZIAMENTO A QUANTI HANNO COLLABORATO PER REALIZZARE QUESTO IMPORTANTE CONVEGNO, ALLE AUTORITÀ INTERVENUTE, AGLI AMICI RELATORI E A TUTTI QUANTI SONO QUI, SPINTI DA CURIOSITÀ E INTERESSE, A SEGUIRE I LORO INTERVENTI.

- MA LA MIA GRATITUDINE NON SAREBBE COMPIUTA, SE NON DEDICASSI QUALCHE CENNO ANCHE A QUELLE FIGURE SEMISCONOSCIUTE E SPESSO TRASCURATE CHE PERÒ HANNO RESO POSSIBILE QUESTA NOSTRA MANIFESTAZIONE: SENZA DI LORO NON SAREMMO QUI A PARLARE DI BITCOIN! ALLUDO ALLE…

FUNZIONI DI HASH

grazie!

Notate che non sono ‘funzioni di Hash’, come le funzioni di Eulero o di altri famosi matematici… : non c’è un prof. Hash!

Sono proprio ‘funzioni di hash’, così denominate scherzosamente per la loro caratteristica di distruggere tutto quanto viene dato loro in pasto. In italiano si potrebbero rendere con ‘funzioni polpetta’.

Ora, non intendo certo rubare la palla ai relatori successivi, ma vorrei soltanto giustificare il mio ringraziamento iniziale con una brevissima esposizione delle loro caratteristiche funzionali e della loro effettiva indispensabilità nella costruzione di moltissime strutture crittografiche, compresi appunto i nostri bitcoin.

- Ho già detto che queste funzioni sono distruttive, e in effetti ingeriscono tutti i bit che vengono passati alla loro competenza (*input*) , rimescolandoli e metabolizzandoli fino ad ottenere come prodotto finale un'uscita (*output*) di pochi byte di una lunghezza fissa e predefinita, che non a caso viene chiamata '*digest*' (in italiano di solito 'impronta')

- E’ singolare che queste funzioni siano state inizialmente considerate del tutto inutili e insuscettibili di applicazioni concrete, destinate ad interessare soltanto la matematica pura...
- In realtà la loro attività preziosa e insostituibile è quella di catalogare ogni possibile sequenza di bit con un numero di protocollo **unico**, l’impronta,appunto.

- Perché l’impronta è sempre e solo un numero, anche se gli informatici preferiscono rappresentarlo con cifre in base esadecimale (16), anziché con la nostra consueta notazione decimale. Una minima variazione nell’ ingresso, anche di un solo bit, genererà un’impronta del tutto differente, consentendo così un controllo efficace e rigorosissimo dell’integrità dei messaggi.

- Ad esempio l'impronta della funzione di hash SHA256 calcolata sulla parola 'ghislieri' è:

2e82d291786bc0bd23b2863da340b9dba30e4211c6f984e445f09e554ff77723,
un numero che può essere 'tranquillamente' rappresentato in decimale come
938.518.644.378.977.054.908.393.685.334.990.402.413.286.836.893.475

- Se però do in pasto alla stessa funzione SHA256 la parola 'Ghislieri', che differisce dalla precedente di un solo bit, la polpetta risultante dal metabolismo di hash sarà:

773ee7bd1575d9cba2c0a082c1f34d860fea88b86552f09bdbeee4d09f8d285d,
chiaramente del tutto differente, e sempre rappresentabile in decimale come
10.385.106.080.308.036.956.742.695.503.501.382.700.860.037

- Sono numeri che appaiono enormi (tranne che per i crittografi...), ma d'altra parte come sarebbe possibile diversamente assegnare un numero di protocollo unico a tutte le combinazioni possibili, dalla lettera singola all'intera Enciclopedia Treccani? E non parliamo solo di testi, ma di qualunque sequenza di bit, visuale, sonora ecc.

- E allora questa caratteristica unica di fungere da ‘cane da guardia’, di segnalare immediatamente qualunque variazione, anche piccolissima, in una sequenza binaria, risulta preziosa laddove è essenziale garantire l’integrità dei dati testuali, come nella firma elettronica.
- Ma è addirittura insostituibile nella costruzione di quella catena di dati relativi alle transazioni operate in bitcoin, la *blockchain* (le cui caratteristiche verranno ampiamente illustrate dai relatori successivi): sostanzialmente un database che contiene la registrazione di TUTTE le operazioni effettuate con i bitcoin a partire dal gennaio 2009, strutturato appunto come una catena in cui le funzioni di hash validano di volta in volta gli anelli precedenti e garantiscono il sicuro insuccesso di chi intendesse modificare anche soltanto un bit di una qualsiasi transazione.

- Un database accessibile al controllo di tutti, naturalmente di ampiezza costantemente crescente col susseguirsi delle transazioni (ad oggi navighiamo sui 58 GB...).
- Appare incredibile come questa costruzione informatica *open source* sappia coniugare la trasparenza di tutte le operazioni, la (relativa) anonimità delle stesse con la garanzia dell'assenza di manipolazioni che possano attentare all'integrità dei dati. Il tutto senza l'intervento di alcuna organizzazione centrale finanziaria, ma gestito da una struttura protocollare distribuita di *peer* su base sostanzialmente volontaria. Un fenomeno veramente curioso ed interessante, che vale la pena di conoscere meglio per approfondirne le caratteristiche e, aggiungo io, anche per riconoscere i meriti delle umili e, spero ora un po' meno misconosciute, **FUNZIONI DI HASH!**