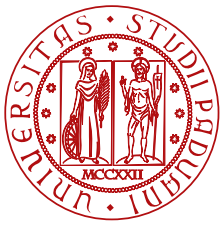


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA INFORMATICA

Blockchain e Learning Analytics: Una Nuova Frontiera per la Gestione dei Dati

Relatore
Prof. Varagnolo Damiano

Laureando
Speranza Ludovico

Correlatore
Cognome Nome

ANNO ACCADEMICO 2024-2025

Data di laurea GG/MM/AAAA

Apes. Together. Strong.

Sommario

Descrizione dell’obiettivo della tesi.

Indice

Elenco delle figure

Capitolo 1

Blockchain

Nonostante il grande successo della trasformazione digitale, molti ambiti delle nostre interazioni sono rimasti indietro e non si sono adattati al digitale; un esempio su tutti è quello del settore legale. Inoltre, la digitalizzazione presenta anche lati oscuri [1]: i nostri dati sono controllati da organizzazioni il cui obiettivo principale è il profitto, gli hacker tentano continuamente di sottrarre informazioni personali, e gli Stati monitorano le attività dei cittadini per mantenere l'ordine e il controllo. Tutto ha origine nel 2008, con la pubblicazione online dell'articolo "Bitcoin: A Peer-to-Peer Electronic Cash System" da parte di Satoshi Nakamoto, pseudonimo che potrebbe celare un individuo o un gruppo di persone. L'obiettivo del documento era quello di creare un sistema puramente peer-to-peer per il trasferimento di valore tra le parti, eliminando la necessità di intermediari come banche o istituti finanziari. Il concetto chiave della blockchain è la possibilità di creare una generazione di piattaforme decentralizzate e disintermediate, nelle quali la fiducia tra le parti non è garantita da un'organizzazione centrale, ma dalla piattaforma stessa, attraverso un meccanismo algoritmico essa si riferisce, in senso più ampio, all'insieme di tecnologie che rendono possibile la decentralizzazione, basandosi sul modello peer-to-peer, in sostanza consente di immaginare un mondo organizzato in modo completamente diverso. La blockchain è un database completamente decentralizzato, di cui ogni nodo possiede una replica e concordano, sulla base di un algoritmo comune, un'unica versione aggiornata delle informazioni. Si tratta di una struttura dati nella quale è possibile solo aggiungere nuovi dati, senza possibilità di cancellare quelli precedenti, garantendo così uno storico completo delle modifiche. Per chiarire meglio questo concetto, possiamo immaginare la blockchain come una barriera corallina: lo strato attivo corrisponde alla parte più recente, che viene continuamente modifi-

Figura 1.1: Statua raffigurante Satoshi Nakamoto

cata con nuove aggiunte, mentre gli strati più vecchi rimangono immutabili e accessibili solo per attività di consultazione. Può essere quindi rappresentata come una lista in continua crescita di "blocchi" collegati tra loro e protetti mediante crittografia. Oltre alla decentralizzazione e all'immutabilità, il terzo grande punto di forza è il consenso: nuove transazioni possono essere registrate solo quando la maggioranza dei partecipanti alla rete dà il proprio consenso.
I blocchi sopra citati sono composti da 7 campi principali:

Block: rappresenta il numero del blocco all'interno della blockchain. Ogni volta che un nuovo blocco viene aggiunto, questo numero aumenta progressivamente.

Nonce: è un valore, spesso composto da 32 bit (come nel caso di Bitcoin), fondamentale perché viene cercato dai miner per generare un hash valido che soddisfi i requisiti di difficoltà della chain. Serve a modificare l'input del calcolo dell'hash e viene aggiornato continuamente durante il processo di mining.

Timestamp: indica il momento esatto in cui il blocco è stato aggiunto alla blockchain.

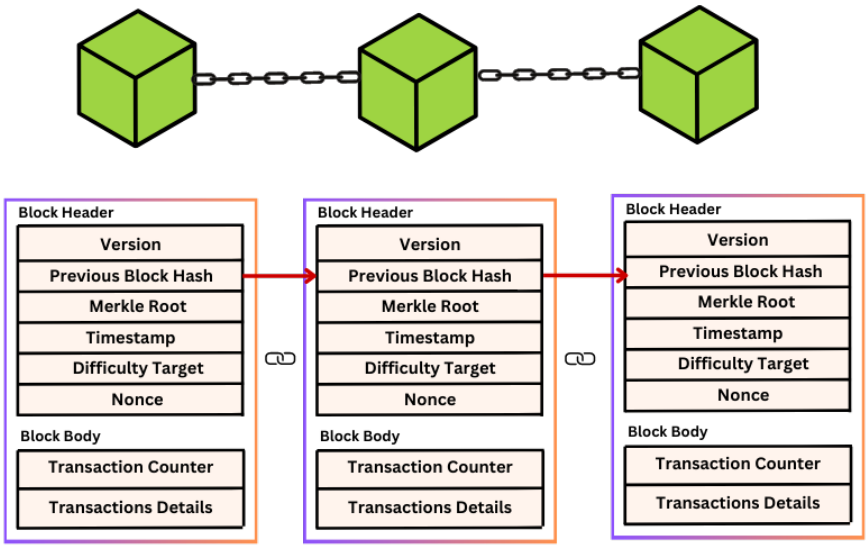
Transaction: contiene o un valore (ad esempio un importo monetario), o il corpo del messaggio, o uno smart contract.

Transaction n: si riferisce a una specifica transazione tra le diverse transazioni presenti all'interno del blocco.

Prev Hash: è l'hash del blocco precedente, che funge da chiave di collegamento tra i blocchi. Questo meccanismo garantisce che ogni blocco sia dipendente dal precedente, creando una catena ininterrotta.

Hash: rappresenta l'impronta digitale del blocco. È un valore alfanumerico unico calcolato applicando una funzione hash. L'hash è deterministico, nel senso che lo stesso input produrrà sempre lo stesso hash; tuttavia, non è possibile risalire ai dati originali partendo dall'hash. Ogni minima modifica ai dati genera un hash completamente diverso, garantendo così l'integrità e l'immutabilità delle informazioni.

Figura 1.2: Esempio di catena



1.1 Blokchain 1.0

1.1.1 Bitcoin

L'aspetto rivoluzionario di Bitcoin è la creazione di una forma monetaria completamente decentralizzata. La sua rete è composta da una serie di componenti chiamate nodi, che collaborano tra loro per raggiungere il consenso su una sequenza di transazioni. Essendo basato sulla blockchain, Bitcoin è completamente decentralizzato: non è gestito da un'entità centrale, come avviene per esempio con le banche. Questo comporta un vantaggio significativo, ovvero una maggiore resistenza a comportamenti anomali o malintenzionati. In un sistema centralizzato, infatti, il nodo centrale rappresenta un punto di vulnerabilità che potrebbe compromettere l'intero sistema. Un'altra caratteristica fondamentale di Bitcoin è la disintermediazione: ogni transazione coinvolge direttamente il mittente e il destinatario, che raggiungono il consenso tramite l'approvazione della rete. Bitcoin è quindi costituito da una rete di sistemi indipendenti, i nodi, capaci di mantenere un registro univoco e condiviso delle transazioni. Questo sistema garantisce che un singolo bitcoin non possa essere speso più di una volta.
In tal modo, Bitcoin soddisfa le funzioni critiche di ogni valuta:

- Immagazzinare valore.
- Essere un'entità contabile.
- Trasferire valore tra le parti.

La sua blockchain è di tipo permissionless, il che significa che chiunque può eseguire il software del nodo e collegarsi alla rete senza alcuna forma di autenticazione. In altre parole, chiunque può unirsi alla rete senza necessità di approvazioni o permessi da parte di un'autorità centrale. Un elemento fondamentale è il wallet che consente agli utenti di gestire i propri bitcoin, conservati in veri e propri conti. Ogni bitcoin è associato a un address, protetto da una coppia di chiavi crittografiche: La chiave pubblica rappresenta l'indirizzo del conto ed è accessibile a tutti. La chiave privata rappresenta l'identità del proprietario ed è necessaria per accedere e autorizzare le transazioni. Bitcoin utilizza la crittografia asimmetrica, un sistema inventato nel 1976 da Whitfield Diffie e Martin Hellman. Questo metodo permette di cifrare un messaggio con una chiave (pubblica) che può essere decifrato solo con l'altra chiave (privata), e viceversa, ciò determina due vantaggi principali. Il primo è Hashing dei dati che trasforma qualsiasi quantità di dati in una stringa cifrata di dimensioni fisse, chiamata digest.

L'hashing è:

Deterministico: ovvero lo stesso input produce sempre lo stesso digest.

Lunghezza fissa: il digest ha una lunghezza predeterminata, indipendentemente dalla dimensione dell'input.

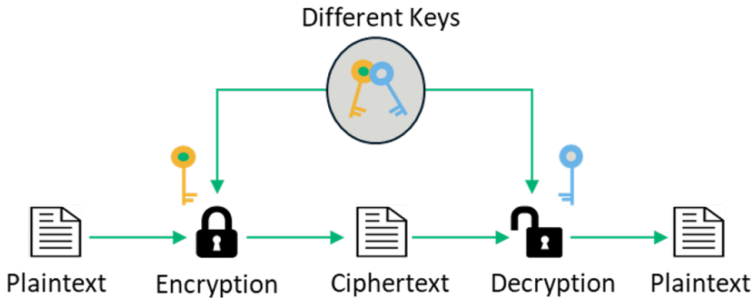
Unico: è estremamente improbabile che due input diversi generino lo stesso digest.

Non invertibile: è impossibile risalire all'input originale partendo dal digest.

Instabile: una minima modifica all'input genera un digest completamente diverso.

Il secondo grosso vantaggio è la firma digitale che garantisce unicità e non ripudiabilità. Quando un utente X invia a un utente Y un messaggio firmato digitalmente, X crea l'hash del messaggio e lo cifra con la propria chiave privata. Y, ricevendo il messaggio, può verificare l'hash con la chiave pubblica di X, garantendo così l'integrità del messaggio e la validità del mittente.

Figura 1.3: Esempio di cifratura asimettrica



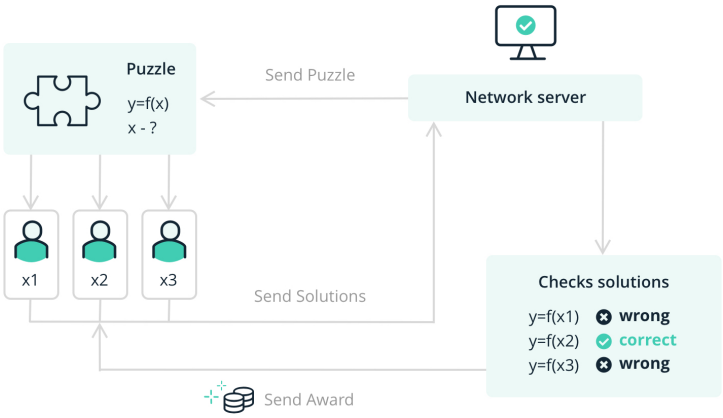
Bitcoin è open source, il suo codice sorgente è pubblico e rilasciato sotto licenza MIT, che permette la massima libertà di utilizzo per progetti futuri. Nonostante il grande successo riscosso, Bitcoin presenta diverse criticità che ne limitano l'efficacia e l'adozione su larga scala. Una delle principali problematiche riguarda la scalabilità: l'algoritmo Proof of Work (PoW) su cui si basa il sistema, pur garantendo sicurezza, tende a limitare l'efficienza complessiva, rendendo difficile gestire un alto numero di transazioni in tempi brevi. Un altro aspetto delicato è quello relativo a privacy e visibilità. Sebbene gli account su Bitcoin siano pseudonimi, tutte le transazioni effettuate sulla rete sono pubblicamente visibili. Questo permette, tramite un'analisi approfondita dei movimenti della valuta, di risalire potenzialmente all'identità dei possessori, mettendo in discussione la reale riservatezza del sistema. La difficoltà di aggiornamento rappresenta un'ulteriore sfida. Il complesso sistema di governance, che coinvolge minatori, sviluppatori e utenti, rende particolarmente arduo raggiungere un consenso per apportare modifiche al protocollo.

Ciò può rallentare l'implementazione di miglioramenti necessari. Infine, vi è il problema delle commissioni elevate, che si acuisce nei periodi di intenso utilizzo della rete. In queste circostanze, effettuare transazioni rapide può diventare molto costoso, specialmente se il tasso di cambio tra Bitcoin ed Euro è particolarmente alto.

1.1.2 Proof Of Work

Come accennato in precedenza, la sicurezza e la validità delle transazioni in Bitcoin sono garantite da un algoritmo di consenso chiamato Proof of Work (PoW) [2]. Un algoritmo di consenso è un protocollo condiviso da tutti i nodi della rete, che consente loro di concordare su una visione unica delle transazioni. In primo luogo, ogni nodo condivide con gli altri nodi della rete le transazioni ricevute nell'ultimo intervallo di tempo, firmate digitalmente. Successivamente, ciascun nodo costruisce in memoria un elenco delle transazioni ricevute dagli altri nodi con cui è in comunicazione. Queste transazioni vengono ordinate, e viene verificato che ogni account disponga dei fondi necessari per completare i pagamenti. A questo punto, i nodi partecipano a una competizione crittografica per individuare, in modo casuale, il nodo che avrà il diritto di pubblicare il blocco successivo. Quando un nodo vince la competizione, individua un blocco valido e lo comunica a tutti gli altri nodi. Questi ultimi verificano il blocco e lo aggiungono alla propria copia della blockchain.

Figura 1.4: Esempio di funzionamento del PoW



Nell'ambito blockchain si sente spesso parlare di fork [3], ma non è sempre chiaro cosa si intenda con fork e cosa effettivamente comporti. Sebbene il termine venga spesso utilizzato per indicare la divisione di una blockchain in realtà esso racchiude un insieme di diversi possibili scenari. Una fork è una situazione in cui accade una delle seguenti cose:

- Può accadere che due o più nodi trovino contemporaneamente una soluzione valida. In tal caso, si generano due blocchi con lo stesso genitore, creando una biforcazione nella catena, detta fork. Quando viene individuato il blocco successivo, l'algoritmo seleziona la catena più lunga (in termini di difficoltà complessiva) come valida. Questo caso viene denominato come fork regolare.

- Le regole della Blockchain sono cambiate in maniera retrocompatibile e tutti i nodi condividono la stessa cronologia delle transazioni, in questo caso si tratta di soft fork e non c'è una divisione della blockchain.

- Le regole della Blockchain sono cambiate in maniera non retrocompatibile, ma tutti i nodi si aggiornano alle nuove regole e condividono la stessa cronologia delle transazioni, anche in questo caso non si ha una divisione della blockchain, ma si tratta di un hard fork.

- Le regole della blockchain sono cambiate in maniera non retrocompatibile e nodi diversi hanno opinioni diverse sulle regole della blolckchain, non condividendo la stessa cronologia delle transazioni, abbiamo quindi un hard fork con chain split, perchè avviene una divisione della chain e il consenso sulla cronologia delle transazioni e perso definitivamente.

Questo criterio, basato sul "principio di maggior lavoro", permette a tutti i nodi di concordare su una versione univoca della blockchain e di risolvere le biforcazioni. Un blocco viene considerato finalizzato quando sono stati generati almeno sei blocchi successivi. Questo livello di profondità rende economicamente insostenibile, per una minoranza di nodi malintenzionati, modificare i dati della blockchain, poiché richiederebbe un effort computazionale estremamente elevato pari al 51% della potenza totale di calcolo del network e anche se ipoteticamente un miner riuscisse a raggiungere tale potenza non sarebbe comunque in grado di modificare le vecchie transazioni, poiché dovrebbe ricalcolare la PoW di tutti i blocchi successivi, mentre gli altri miner onesti continuano a minare sulla blockchain corretta. Un attacco di questo tipo richiederebbe l'utilizzo di una quantità incredibile di risorse per l'attacker. Se qualcuno effettivamente riuscisse a mettere insieme più del 51% della potenza di calcolo, sarebbe molto più redditizio per lui seguire le regole della blockchain. Nonostante i numerosi vantaggi che il Proof of Work offre, questo algoritmo di consenso presenta alcune significative criticità. La principale è il massiccio consumo di energia, un aspetto che, paradossalmente, costituisce anche uno dei suoi punti di forza, rendendo estremamente costoso e complesso attaccare la rete. A questo si aggiunge la scarsa scalabilità del sistema, che si traduce in una certa lentezza nell'elaborazione delle transazioni e nell'aumento delle commissioni, soprattutto nei periodi di maggiore attività sulla rete. Infine, il Proof of Work tende a creare una sorta di discriminazione geografica: attualmente, la mag-

gior parte dei miner si concentra in aree dove il costo dell'elettricità è più basso, limitando la partecipazione globale e accentuando disparità economiche e infrastrutturali.

1.2 Blokchain 2.0

1.2.1 Ethereum e Proof Of Stake

Bitcoin, per sua natura, non è stato concepito come un ambiente di sviluppo e offre funzionalità di programmazione molto limitate, mentre la community ha sempre voluto mantenere la piattaforma focalizzata sullo scambio di valori, lasciando spazio a progetti alternativi che miravano ad ampliare le possibilità della tecnologia blockchain. È in questo scenario che nasce Ethereum, la prima blockchain progettata specificamente per supportare lo sviluppo di applicazioni decentralizzate. Co-fondata nel 2013 da Vitalik Buterin, un programmatore russo-canadese che in passato aveva collaborato con Bitcoin Magazine, Ethereum si distingue da Bitcoin per molteplici innovazioni. Una delle principali novità introdotte è la possibilità di utilizzare due tipi di conti: i tradizionali conti posseduti tramite una coppia di chiavi pubblica e privata, come in Bitcoin, e i conti associati agli Smart Contract [2]. Gli Smart Contract, concettualizzati per la prima volta da Nick Szabo nel 1994, sono programmi in grado di eseguire automaticamente azioni predeterminate una volta soddisfatte certe condizioni. Questi contratti decentralizzati garantiscono che le regole siano rispettate senza la necessità di intermediari, eliminando possibili interferenze. Un esempio pratico di utilizzo degli Smart Contract è il crowdfunding. Un utente potrebbe pubblicare un progetto stabilendo un obiettivo economico e un tempo limite per raggiungerlo. Lo Smart Contract, in questo caso, si occuperebbe di raccogliere i fondi dagli investitori e trattenerli fino al completamento della campagna. Se l'obiettivo viene raggiunto, i fondi vengono trasferiti automaticamente al creatore del progetto; in caso contrario, tornano ai donatori. Questa struttura elimina la necessità di intermediari e rende il processo completamente trasparente. Ethereum utilizza Solidity, un linguaggio di programmazione simile a JavaScript, per scrivere Smart Contract. La flessibilità offerta da questa piattaforma ha rivoluzionato il mondo delle blockchain, permettendo la creazione di applicazioni decentralizzate e inaugurando l'era delle cosiddette blockchain 2.0. Inizialmente veniva utilizzato il meccanismo di consenso Proof of Work (PoW), lo stesso di Bitcoin, però per migliorare efficienza e sostenibilità, il 15 settembre 2022, con l'aggiornamento noto come The Merge, Ethereum è passato al Proof of Stake (PoS). Questo cambiamento ha segnato un cambio epocale per la piattaforma ed ha portato con se enormi vantaggi. Il PoS riduce drasticamente il consumo energetico, abbassandolo del 99,9% rispetto al PoW, ed elimina la necessità di hardware costoso per il mining, rendendo la partecipazione al consenso più accessibile. Con l'introduzione del PoS, Ethereum ha reso possibile lo staking, un processo in cui i partecipanti bloccano una quantità di criptovaluta in uno Smart

Contract per sostenere la sicurezza e il funzionamento della blockchain. Attraverso lo staking, gli utenti possono diventare validatori e partecipare attivamente alla convalida delle transazioni e alla creazione di nuovi blocchi. Lo staking offre numerosi vantaggi: aumenta la sicurezza del sistema rendendo gli attacchi costosi e difficili da eseguire, favorisce la decentralizzazione permettendo a più utenti di contribuire alla rete e riduce il consumo energetico rispetto al mining tradizionale. I partecipanti allo staking ricevono ricompense proporzionali alla loro attività, sotto forma di nuove criptovalute o una parte delle commissioni generate dalle transazioni. Le ricompense non si limitano ai validatori che propongono nuovi blocchi, ma includono anche coloro che verificano e confermano la validità dei blocchi proposti da altri. Questo sistema rende lo staking una delle principali motivazioni per partecipare al consenso su Ethereum, favorendo un equilibrio tra efficienza e incentivi economici. Inoltre, il sistema garantisce una maggiore scalabilità e permette ai validatori onesti di proteggere la rete da attacchi, penalizzando economicamente i nodi malevoli o inattivi per almeno il 50% del tempo attraverso un meccanismo chiamato slashing che risulterà totale per i primi andando a sottrarre in maniera permanente le risorse depositate nello stake, mentre sarà parziale per i secondi che si vedranno sottratti solo i guadagni ricevuti ma non saranno rimossi dalla chain. Nonostante questi vantaggi, il PoS presenta alcune criticità. Essendo una tecnologia più recente rispetto al PoW, è meno testata e potrebbe essere più vulnerabile a eventi imprevisti, i cosiddetti Black Swan. Inoltre, il fatto che il controllo della rete sia legato al possesso di criptovalute pone un rischio di centralizzazione: chi dispone di maggiori capitali potrebbe accumulare una quantità significativa di token, aumentando la propria influenza sul sistema.

Capitolo 2

Sicurezza

2.1 Benefici in termini di privacy e sicurezza

2.2 Introduzione al web 3.0

L'idea del Web 3.0 è stata inzialmente utilizzata in stretta connessione con il concetto di web semantico, terminologia creata da Berners Lee in un articolo di *Scientific American* del 2001 per descrivere un nuovo web. La sua idea è che, così come il web 2.0 permette di collegare pagine web, a livello di visualizzazione, il web semantico deve permettere non solo di collegare pagine tra loro ma anche i dati contenuti in esse [4]. La visone del web semantico è una visione di dati interconnessi e navigabili che possono essere usati da chiunque. L'esigenza di socializzazione dei dati è ancora più importante oggi che l'intelligenza artificiale può costruire modelli a partire dai dati grezzi, sulla base di algoritmi generali. L'obiettivo del web 3.0, riconosciuto da Gavin Wood, co fondatore di Ethereum, è quello di creare un web decentralizzato, dove i dati sono posseduti dagli utenti e non da poche aziende. Sta nascendo la consapevolezza di re-decentralizzare i servizi web. Questa idea è diffusa da molto tempo, ma adesso finalmente esistono le tencologie che permettono di realizzarla, come la blockchain. Così si potrebbero ottenere numerosi vantaggi, tra i quali:

Decentralizzazione: Non è necessario alcun permesso da parte di un'autorità centrale per caricare qualcosa sul web. Questo fornisce una protezione contro qualsiasi forma di censura e controllo. Il web tornerebbe ad essere un sistema neutrale.

izzazione degli accessi: E' possibile offrire un accesso a chiunque abbia una connessione a internet, senza discriminazioni su età, sesso, razza, religione e posizione geografica.

Uptime dei servizi: Non essendoci nodi centrali, non esiste un punto di fallimento. Se un nodo va giù, il servizio è comunque disponibile.

Possesso dei dati: Gli utenti riprenderebbero possesso dei propri dati potendo decidere con chi condividerli e in che modo, potendo potenzialmente guadagnare dalla vendita dei propri dati alle grandi multinazionali come FaceBook e Google le quali hanno tantissime informazioni sugli utenti e gli advertiser che pubblicano le pubblicità sulle loro piattaforme pagano milioni per avere questi dati.

sistenza dei dati: I dati non possono essere cancellati, a meno che non venga cancellata l'intera blockchain, in quanto verranno salvati in maniera ridondante su diversi nodi distribuiti indipendentemente.

Bibliografia

[1] klaus Schwab, *La quarta rivoluzione industriale*. FrancoAngeli, 2019, ISBN: 9788891743008.

[2] N. Attico, *Blockchain, guida all'ecosistema*. Guerini Next, 2018, ISBN: 9788868962180.

[3] R. B. Gianluca Chiap Jacopo Ranalli, *Blockchain, tecnologia e applicazioni per il business*. Hoepli, 2019, ISBN: 9788820389253.

[4] TED, *Tim Berners-Lee on the Next Web*, Accessed: 2023-10-01, 2009. indirizzo: `https://www.youtube.com/watch?v=OM6XIICm_qo&ab_channel=TED`.