

QUANTUM COMPUTING PER L'OTTIMIZZAZIONE DELLE RETI MOBILI (4.5G E 5G)

Andrea Boella, Michele Ludovico,
Giuseppe Minerva, Mauro Alberto Rossotto

Il quantum computing è tra le tecnologie innovative di maggior interesse, accompagnato da un "hype" sui mezzi di comunicazione per le potenzialità che sembrano promettere un impatto disruptive in diversi campi.

La competizione per la leadership tecnologica nel quantum computing è partita da alcuni anni e l'interesse è cresciuto negli ultimi tempi grazie ai progressi della ricerca e al recente annuncio sul raggiungimento della "quantum supremacy" da parte di Google, cioè l'individuazione di un primo esempio reale di algoritmo implementato sul Quantum Computer, che non può essere elaborato in tempi di calcolo accettabili sui computer classici.

Questi risultati hanno creato grandi aspettative, ma, allo

stato attuale di sviluppo della tecnologia, si prevede che la disponibilità "commerciale" di quantum computers con potenze di calcolo elevate e la loro applicazione su larga scala, si verificherà su un orizzonte temporale stimabile tra i 5 e i 10 anni.

Ciò non toglie che il quantum computing possa già essere applicato ad alcuni casi d'uso, sfruttandone opportunamente le capacità computazionali attualmente disponibili. È quanto è avvenuto nel contesto TIM, in cui è stato applicato ad un problema di pianificazione delle reti mobili 4.5G e 5G. I risultati sono incoraggianti avendo ottenuto una maggiore rapidità di esecuzione (con un fattore 10x) e una qualità media superiore rispetto ai metodi tradizionali di ottimizzazione.

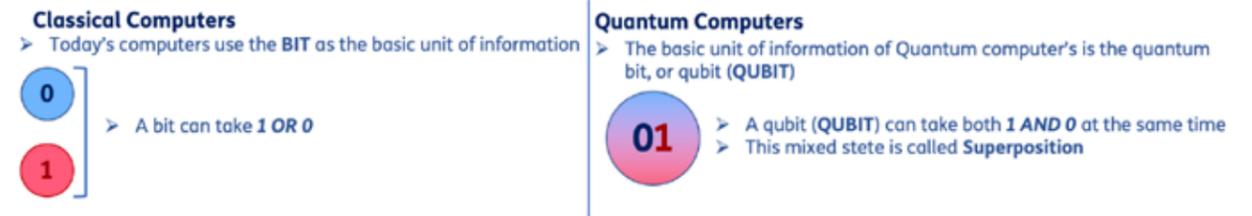
Introduzione al quantum

L'idea che ha dato origine allo sviluppo del quantum computing nasce dall'osservazione che i fenomeni della natura e le aree scientifiche che dal loro studio sono state derivate (fisica, chimica, biologia...), sono governati dai fenomeni della meccanica quantistica. Pertanto, per progredire ulteriormente nella conoscenza di questi campi, è necessario disporre di un computer che abbia una logica computazionale che riproduca gli stessi principi, cioè - in altri termini - un'analogia modale di funzionamento. In questo modo, tutte le attività di ricerca e sviluppo (elaborazione di dati, definizione di modelli, simulazioni, analisi predittive, ecc.) potranno essere effettuate più facilmente e in maniera più accurata, perché più precisa potrà essere la rappresentazione dei fenomeni da analizzare. La caratteristica di questi problemi è quella di avere una complessità computazionale che cresce esponenzialmente all'aumentare delle dimensioni dei dati di input. La complessità di questi problemi, dovuta alla numerosità delle variabili, impedisce ai computer classici di eseguire simulazioni o ricerche di ottimizzazione efficaci a causa del loro approccio "lineare e sequenziale". La crescita della capacità computazionale degli elaboratori classici non è una soluzione praticabile in questo ambito, per cui si è guardato ai fenomeni della meccanica quan-

tistica per realizzare degli elaboratori che cambiasse l'approccio e la metodologia di lavoro. Sono stati così teorizzati e poi progettati i primi computer quantistici che, grazie ad alcune caratteristiche peculiari della meccanica quantistica, permettono un'elaborazione parallela e probabilistica in grado di affrontare e risolvere questi casi complessi.[1] Questo nuovo approccio "quantistico" è generalizzabile, perché questa tipologia di problemi si incontra in vari contesti come la chimica dei materiali, la meteorologia, la finanza. Pur disponendo di una potenza di calcolo elevata, i quantum computers non sostituiranno i computer classici, ma lavoreranno in sinergia con questi secondo un modello di computazione ibrida classica - quantistica. Le due tecnologie, infatti, sono adatte a classi distinte di problemi: alle famiglie di problemi elaborabili con computer classici, si aggiungono quindi le famiglie adatte ai quantum computers. Nei computer classici, l'unità di informazione attraverso cui sono codificati i dati da processare è rappresentata dal bit che può assumere due valori in maniera fra loro esclusiva (0 oppure 1). Nei computer quantistici l'unità di informazione è, invece, il qubit (quantum bit) che, grazie ai principi della meccanica quantistica che ne sottendono l'implementazione, può assumere non solamente i valori 0 o 1, ma anche entrambi contemporaneamente (0 e 1), in uno stato di sovrapposizione

intermedio tra i due stati fisici corrispondenti ai valori 0 e 1. Per usare un'analogia, si può pensare ad una moneta che a seconda di come è posizionata può essere rivolta verso il lato "testa" o lato "croce". Se la moneta viene fatta roteare, finché non si ferma entrambe le facce "testa" e "croce" sono contemporaneamente possibili; in questa condizione la moneta è come se si trovasse in uno stato di sovrapposizione. Questo stato, che nel mondo quantum si chiama superposition, conferisce ai quantum computer quelle caratteristiche probabilistiche alla base delle loro potenze di calcolo esponenziali. Le capacità computazionali dei quantum computers si misurano sulla base del numero di qubit che sono in grado di gestire. Nella famiglia dei quantum computer, si possono distinguere alcune tipologie di processori - anche denominati QPU (quantum processing unit) - in base alla tecnologia e alla finalità:

- Quantum gate array: anche denominato universal quantum computer per intendere la macchina in grado di elaborare qualunque processo computazionale complesso (qualunque algoritmo quantistico) in forma massiva (senza limitazioni sulle dimensioni dei dati trattati) in tempi rapidi. Si tratta del quantum computer così come lo immaginiamo per analogia con i computer classici, versatile con prestazioni massime, ma anche molto complesso da realizzare,



1

Confronto tra Classical and Quantum Computing

Il bit, unità di informazione classica, può assumere solo uno dei possibili valori alla volta (0 OR 1) mentre i qubit, unità di informazione quantistica, possono assumere entrambi i valori (0 AND 1) contemporaneamente (superposition)

a causa dell'elevato numero di qubit necessari. Per avere un'idea del grado di interesse attorno al quantum gate array, i vendor che stanno investendo in questa tecnologia sono IBM, Google, Microsoft, Intel, Alibaba. I quantum computer attualmente più potenti supportano circa 50 qubit e per arrivare a disporre di macchine che ne equipaggino alcune centinaia, quantità necessaria per le dimensioni dei problemi che dovrebbero trattare, occorrerà, in base alle attuali previsioni, aspettare alcuni anni (5 - 10). Per coprire questo gap temporale e poter già applicare il quantum computing a casi d'uso reali, la tecnologia si è orientata anche su soluzioni alternative, descritte nei punti seguenti. [2]

- Quantum annealer: è un computer che utilizza una modalità di elaborazione che riproduce un fenomeno fisico specifico: la tendenza naturale che por-

ta un sistema a raggiungere la condizione di equilibrio a minima energia. Modellando quindi un fenomeno ben definito, è un computer adatto a trattare una classe specifica di problemi, quelli di ottimizzazione di natura combinatoria. Il fornitore più rappresentativo della categoria quantum annealer è canadese, D-Wave che fornisce un processore di 2000 qubit e ha in roadmap il rilascio commerciale della sua evoluzione a 5000 qubit nel corso del 2020. La numerosità dei qubit così elevata rispetto al quantum gate array, è dovuta al fatto che i quantum annealer non sono macchine universali, ma ottimali per una specifica famiglia di problemi e nella loro modalità di processing la gestione dei qubit è più semplice, semplificando di conseguenza anche la tecnologia di base. Occorre anche considerare che la potenza di calcolo del quantum annealer non è in rela-

zione diretta col volume di qubit supportati, ma ridotta perché i qubit non sono tra loro connessi a maglia completa.

Attualmente entrambe le tipologie funzionano come dei quantum computer solo se protette in un ambiente pressoché isolato e prossimo alla temperatura dello zero assoluto. Qualunque minimo disturbo (vibrazione, variazione di temperatura, rumore, ecc.) provoca la perdita delle loro proprietà quantistiche e di conseguenza le loro capacità computazionali. Queste tipologie di quantum computer, quindi, devono essere installate in ambienti che garantiscono quelle specifiche condizioni di isolamento tipiche dei laboratori. Per questa ragione l'accesso ai quantum computers attualmente è fornito in modalità cloud attraverso un'interfaccia, tramite la quale si forniscono i dati di input e il codice dell'algoritmo da eseguire scritto in un linguaggio di alto livello e si ottengono i risultati dell'elaborazione.

D-Wave's vision: the future of quantum computing and its applications

Quantum computing has the potential to fundamentally change society and how we interact with technology.

Quantum is poised to disrupt a number of major industries, including transportation, mobility, materials science, medicine, and more, enabling powerful new applications that aren't possible with classical supercomputers.

While it's still early in the lifespan of quantum computing, promising applications with practical business value are emerging.

Cloud access and hybrid quantum/classical computing approaches have made it much easier for virtually any organization or developer to get started building quantum applications.

To date, D-Wave users and customers have built over 200 quantum applications in diverse fields.

Major companies, including Menten AI, Volkswagen, the German Aerospace Center, and the Italian State Railway, have used D-Wave quantum computers to optimize various parts of their R&D, resource allocation, and supply chain.

As quantum computers increase qubit counts and processing power – such as D-Wave's upcoming 5000 qubit Advantage system – applications will become more robust and lead to breakthroughs in critical fields, including drug discovery, transportation, and machine learning, among others.

D-Wave's Quantum Annealing Computers

D-Wave quantum computers use a process called quantum annealing to find solutions to problems expressed in terms of optimizing a certain function. A problem is represented by a graph with real-valued weights on nodes and edges. The graph is mapped onto an arrangement of qubits (nodes) and couplers (edges) inside the quantum processing unit (QPU). During the computation (called an anneal), gradually-evolving forces are applied to the qubit system, to drive it into an energy state that corresponds to an optimal solution to the original problem. Typical anneal times range from 2 to 200 microseconds per input.

A current-generation D-Wave 2000QTM system contains 2000+ qubits with up to six couplers per qubit (fewer on circuit boundaries). This allows the QPU to solve problem graphs with between roughly 64 nodes (dense) and 2000 nodes (sparse), depending on qubit yields.

The QPU operates within a highly shielded environment --- at temperatures below 15 millikelvin, and experiencing less than 50,000x of Earth's magnetic field --- to protect the computation from external noise and improve the probability of a successful outcome. Typical anneal times range from 2 to 200 microseconds per input.

The next-generation AdvantageTM system, to be launched later this year, will contain 5000+ qubits and approximately 15 couplers per qubit, corresponding to problem graphs with between roughly 182 and 5000 nodes.

Dr. Catherine McGeoch
cmcgeoch@dwavesys.com

Per la modalità on-premises, largamente diffusa con i computer classici, occorrerà attendere l'evoluzione tecnologica.

In questa fase transitoria, in attesa della piena maturità del quantum computing, si stanno affacciando sul mercato soluzioni che possiamo definire "QUANTUM INSPIRED". Si tratta di emulatori Hardware e/o Software che permettono di sviluppare codice Quantum Ready, eseguendolo inizialmente su processori lineari come CPU e GPU. Le potenzialità equivalenti di queste soluzioni sono nell'ordine di alcune decine di qubit e quindi comparabili con le potenze di calcolo degli attuali quantum computer di tipo gate array. Rispetto a questi hanno il vantaggio che i loro qubit sono più stabili, visto che utilizzano hardware classico. Ricordando che la potenza di calcolo dei quantum annealer è inferiore rispetto al dato di targa, i quantum inspired sono comparabili - in termini di prestazioni - anche con questa tipologia di quantum computer.

Oltre ai fattori prestazionali, le soluzioni quantum inspired sono interessanti per una serie di ulteriori motivi:

- il software sviluppato per queste piattaforme sarà riutilizzabile sui computer quantistici di potenza elevata, quando si renderanno disponibili
- offrono un ambiente di sviluppo sul quale acquisire esperienza

nella modellizzazione e programmazione quantistica

- la programmazione quantistica degli algoritmi porta già dei primi vantaggi negli scenari di ottimizzazione, poiché la corrispondente modellizzazione è diversa e abilita la ricerca non lineare di soluzioni al problema
- utilizzando processori classici (CPU e GPU), ammettono anche la modalità di installazione on-premises o in Edge (sempre che le loro potenze di calcolo siano sufficienti per le applicazioni su cui si intende impiegarli)

La disponibilità di computer quantici di capacità di calcolo elevate, potrà sicuramente ridurre i tempi di elaborazione di queste stesse procedure sviluppate su quantum inspired e contemporaneamente aiutare a trovare le soluzioni ottime, cioè di maggiore qualità.

Applicazioni del quantum computing

I computer quantistici sono adatti per trattare problemi complessi, che sono rappresentabili con modelli multidimensionali nei quali intervengono un numero elevato di variabili e che richiedono, di conseguenza, tempi di elaborazione molto lunghi, talvolta non compatibili con le tecnologie tradizionali.

I casi d'uso che presentano queste caratteristiche sono tipicamente:

- problemi di ottimizzazione
 - ottimizzazione nell'assegnazione/riuso di risorse soddisfacendo certi vincoli, come per esempio nei casi assimilabili al problema di colorazione delle mappe
 - ottimizzazione di percorsi (esempio tipico è il travel salesman problem) per gestione di veicoli, robot, droni, aerei, flusso di traffico ecc.
- ricerche esaustive su spazi di ricerca di grosse dimensioni
 - applicazioni legate alla crittografia, per sviluppare nuovi protocolli di cifratura resistenti ai quantum computer
- classificazione di dati di grosse dimensioni o per identificazione di pattern ricorrenti
 - applicazioni per il Machine Learning di image processing, analisi predittive, diagnosi
 - set ottimale dei pesi di una rete neurale.

Le applicazioni descritte sono non solo complesse a livello computazionale, ma anche - nella maggior parte dei casi - time-critical.

La gestione real-time di questi casi d'uso, che tipicamente viene soddisfatta collocando la logica di controllo in prossimità del punto di attuazione, al momento non è ancora possibile con i quantum computers ma lo diventerà col miglioramento della tecnologia.

Nella prospettiva che il quantum computing diventi una tecnologia adatta per un'installazione on-premises, si può pensare ad un suo utilizzo a livello di edge, proprio per la gestione delle applicazioni time-critical.

Un'alternativa, per disporre di capacità computazionali quantistiche all'edge, è rappresentata dalle soluzioni quantum inspired, presumibilmente realizzabili in tempi più brevi e compatibilmente con le prestazioni che sono in grado di garantire [3].

Un esempio di ottimizzazione della rete mobile: la pianificazione degli identificativi di cella (PCI)

Nell'ambito delle telecomunicazioni, problemi di ottimizzazione sono presenti in differenti contesti in cui le risorse sono limitate e condivi-

se. Un primo esempio che è stato affrontato e sperimentato con le tecniche di quantum computing è la pianificazione degli identificativi di cella (PCI: Physical Cell Identifier) nelle reti mobili 4.5G e 5G. I

n ambito radiomobile, si definisce "cella" l'unità elementare di territorio corrispondente ad una antenna trasmittente e ad una specifica banda frequenziale. L'identificazione della cella, proprio attraverso il PCI, consente ai terminali mobili di gestire le procedure di mobilità nel passaggio tra celle geograficamente adiacenti.

L'identificativo PCI è assegnato a ciascuna cella della rete mobile all'interno di un set predefinito di valori, variabile da sistema a sistema in base alle specifiche caratteristiche della rete. In termini generali, il PCI è composto da due differenti campi, denominati - rispettivamente - Group_ID e Cell_ID; il valore risultante dell'identificativo è espresso, in funzione delle

due componenti, mediante la relazione:

$$[Eq1] \quad PCI=3 \cdot Group_ID + Cell_ID$$

La coppia Group_ID e Cell_ID è usata sia per la generazione dei canali di sincronizzazione utilizzati dal terminale mobile per agganciarsi correttamente alla cella sulla quale intende operare, sia per generare il segnale di riferimento della cella medesima (un vero e proprio "faro") la cui individuazione è necessaria affinché il terminale possa decodificare le informazioni di servizio trasmesse dal sistema.

L'obiettivo finale della pianificazione dei PCI è la definizione di un "piano" che minimizzi i problemi correlati ai cosiddetti casi di "collision" (PCI uguali assegnati a celle adiacenti) e "confusion" (PCI uguali assegnati a celle aventi un'adiacenza in comune) illustrati sinteticamente in Figura 2.

In presenza di tali situazioni, infatti, il terminale mobile può fallire le procedure di mobilità nel passaggio da

una cella all'altra e - di conseguenza - subire una caduta di chiamata. Un ulteriore obiettivo, meno prioritario rispetto al precedente, è evitare l'assegnazione di PCI con lo stesso Group_ID a nodi geograficamente vicini ("adiacenti") e operanti sulla stessa banda frequenziale, condizione utile - insieme all'uso obbligatorio di un medesimo Group_ID all'interno di ciascun nodo - al fine di facilitare le funzionalità di decodifica.

I problemi di "collision" e "confusion" corrispondono, a tutti gli effetti, a veri e propri vincoli di assegnazione (comunemente detti di "riuso"), che discendono dall'esistenza delle adiacenze di rete e/o elettromagnetiche specifiche del sistema 5G o LTE e delle adiacenze di rete dovute alla presenza, nel medesimo territorio, di celle delle reti 2G e/o 3G.

In tale contesto è importante evidenziare come la gestione del passaggio di una connessione in essere dalle reti di precedente generazione alla rete LTE, sia spesso garantita semplicemente definendo "in adiacenza" la cella 2G o 3G di interesse e la banda frequenziale LTE verso la quale si intende far migrare la generica chiamata, senza indicazione di specifiche celle della rete 4.5G.

In termini generali, la definizione di un piano dei PCI che soddisfa il maggior numero possibile di vincoli di "riuso" permette di migliorare le prestazioni della rete di accesso e,

quindi, la qualità del servizio sperimentata dagli utenti.

In particolare, una buona pianificazione consente di ridurre significativamente i tassi di caduta delle connessioni in corso, il tutto grazie alla diminuzione dei fallimenti nell'esecuzione delle procedure di Hand-Over (cioè il passaggio della connessione da una cella ad un'altra). Ciò implica notevoli benefici - in particolare - per le connessioni "voce" (chiamate VoLTE, nel contesto LTE) per le quali la velocità di esecuzione e il corretto completamento della procedura sono di fondamentale importanza per ottenere prestazioni adeguate.

Il numero di differenti PCI disponibili ai fini della pianificazione degli identificativi di cella dipende dalla rete considerata ed è pari a 1008 (con numerazione 0÷1007) nel caso 5G e 504 (con numerazione 0÷503) nel caso LTE. Tali valori sono generati mediante la relazione [Eq 1] in base al Group_ID e al Cell_ID, i cui campi di esistenza sono i seguenti:

per 5G
Group_ID ∈ [0; 355] Cell_ID ∈ [0; 2]

per LTE
Group_ID ∈ [0; 167] Cell_ID ∈ [0; 2]

Al fine di meglio quantificare il potenziale peggioramento delle prestazioni della rete, è associato - al mancato rispetto dei vincoli di "riuso" - un "costo" dipendente sia dalla tipologia del vincolo coinvolto,

sia dalla numerosità delle violazioni introdotte (una generica coppia di celle, infatti, può violare contemporaneamente più vincoli di "riuso", anche della stessa tipologia).

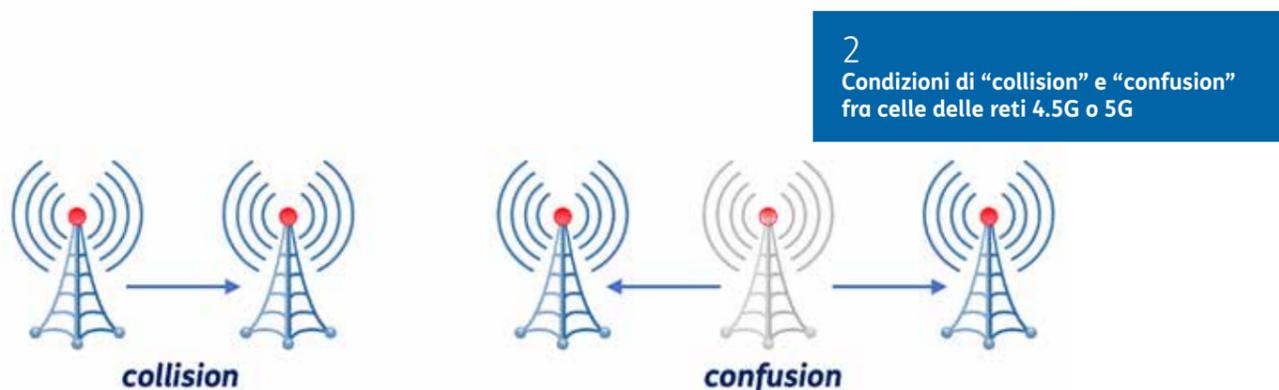
Il costo relativo a ciascun tipo di violazione dipende, quindi, sia dalla conformazione geografica del territorio - attraverso la copertura del segnale radio garantita da ciascuna cella che incide direttamente sulla definizione delle adiacenze - sia dall'importanza che si attribuisce a ciascun vincolo non rispettato, che è propria di ogni Telco Operator poiché derivante dalla lunga esperienza accumulata nel corso dello sviluppo delle reti.

In conseguenza di quanto descritto e data una generica coppia di celle (i,j), i due fattori di costo sono i seguenti:

$C_{i,j}$ costo ("degrado") dovuto all'assegnazione dello stesso PCI alle celle i e j

$S_{i,j}$ costo ("degrado") dovuto all'assegnazione dello stesso gruppo alle celle i e j.

Il costo complessivo (cioè il "degrado" totale) di un piano di assegnazione, è dato da una funzione di costo contenente i contributi (alcuni di valore nullo) dovuti alla potenziale violazione dei vincoli di "riuso" relativi ai PCI e ai Group_ID di ciascuna possibile coppia di celle oggetto della pianificazione.



La funzione assume, quindi, la seguente forma:

$$\text{Tot}_{\text{cost}} = \sum_i \sum_j C_{ij} \cdot v_{ij} + \sum_i \sum_j S_{ij} \cdot w_{ij} \quad [\text{Eq 2}]$$

dove:

$v_{ij} = 1$ se lo stesso PCI è assegnato alle celle i e j

$v_{ij} = 0$ altrimenti

$w_{ij} = 1$ se lo stesso Group_ID è assegnato alle celle i e j

$w_{ij} = 0$ altrimenti.

La pianificazione degli identificativi PCI effettuata in coordinamento con l'algoritmo automatico di individuazione delle adiacenze di rete (operante in closed loop sulla base delle misure raccolte dai terminali) è gestita in TIM attraverso la piattaforma Software Open SON (Self Organizing Network) descritta in [4]. In [5] sono illustrati, inoltre, i risultati ottenuti - in termini di miglioramento delle cadute VoLTE - at-

traverso la pianificazione dei PCI in "closed loop" integrando, all'interno del framework, gli algoritmi di ricerca operativa sviluppati da TIM e impiegati anche nelle fasi di progettazione della rete.

Nel caso della pianificazione dei PCI, in particolare, è utilizzato un algoritmo di tipo "Fast Greedy" [6].

L'approccio alla pianificazione dei PCI basato sul Quantum Computing si inserisce nel contesto Open SON, come descritto in Figura 3.

In particolare, il modulo RAN Orchestrator (corrispondente all'applicativo TIMqual sviluppato "in house" da TIM) si occupa di gestire lo scambio di informazioni con gli elementi di rete, attraverso le API (Application Programming Interfaces) rese disponibili dai sistemi di ACM (Automatic Configuration Management) forniti dai RAN vendors nell'ambito

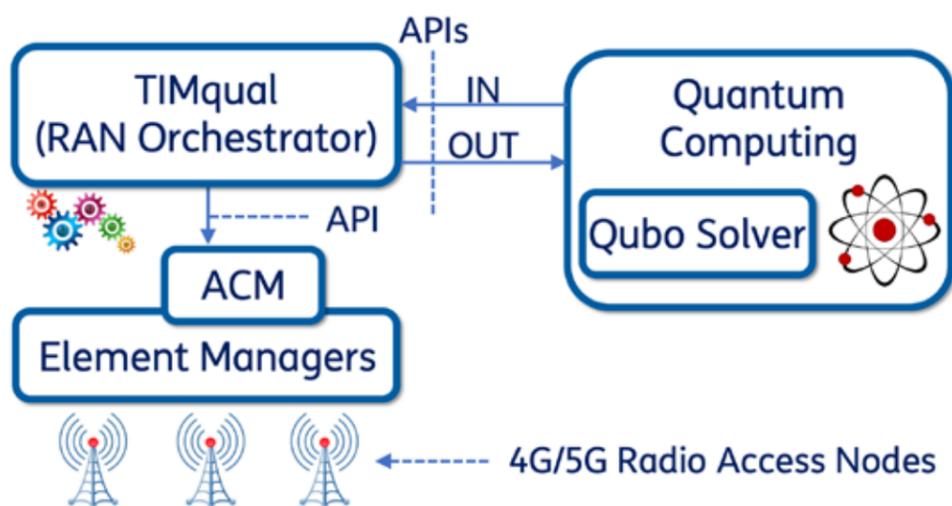
dei sistemi di gestione della rete (Element Managers).

L'applicazione del Quantum Computing è stata pensata con un ulteriore scambio di informazioni verso sistemi esterni (nel caso specifico "hosted" presso D-Wave). Un maggiore livello di integrazione sarà valutato nell'ipotesi di disponibilità di quantum computer gestiti "on premises" all'interno della rete TIM.

Modellizzazione dell'algoritmo di pianificazione PCI in ottica quantum Computing

Modello QUBO

L'attività di pianificazione degli identificativi di cella (PCI), ha l'obiettivo, come descritto, di ottimizzare



REPLY'S VISION: ALGORITMI E CASI D'USO PER IL QUANTUM COMPUTING

Il Quantum Computing è una disciplina al confine tra la fisica, la matematica, l'informatica e l'ingegneria in continuo sviluppo, il cui scopo è quello di svolgere operazioni computazionalmente costose in tempi molto ridotti e, allo stesso tempo, raggiungendo risultati di alta qualità.

La gamma di problemi che può essere affrontata attraverso il formalismo quantistico è molto ampia e spazia dal Machine Learning e Intelligenza Artificiale all'Ottimizzazione Combinatoria.

Nel primo caso si vedono come maggiori protagonisti i cosiddetti Universal Quantum Computer, ovvero dei calcolatori quantistici in cui è possibile accedere direttamente allo stato dei qubit e, pertanto, modificarlo in modo arbitrario. Questo tipo di approccio ha permesso lo sviluppo di algoritmi come il Variational Quantum Eigensolver (VQE) o il Quantum Approximate Optimization Algorithm (QAOA) particolarmente adatte al Machine Learning e alla simulazione di sistemi quantistici. L'idea alla base di queste due strategie consiste nella cooperazione tra macchine classiche e quantistiche: la velocità della Quantum Processing Unit (QPU) viene sfruttata per misurare l'energia del sistema, mentre la parte classica viene sfruttata attraverso un sistema che di fatto impara la strategia migliore per codificare i dati in una macchina quantistica. Di conseguenza diventa possibile trovare, ad esempio, pattern nascosti nei dati per risolvere problemi di Supervised Learning, oppure ottenere il minimo di una funzione arbitraria in un problema di ottimizzazione. È su questo principio che si fondano algoritmi quali Quantum Support

Vector Machines (QSVM) e Quantum Neural Networks (QNN).

Per quanto riguarda i problemi di ottimizzazione combinatoria, risultano adatti i cosiddetti Quantum Annealers. Questi incapsulano un diverso tipo di QPU in cui non è possibile manipolare direttamente lo stato dei qubit, bensì occorre programmare la funzione di energia cui il sistema quantistico è sottoposto. Queste macchine lavorano lasciando che i qubit evolvano naturalmente verso il minimo della funzione d'energia data dal modello di Ising.

L'approccio che fa uso dei Quantum Annealers si presta particolarmente alla soluzione di problemi che difficilmente possono essere riscritti mediante modelli lineari, come, ad esempio, la pianificazione dei Physical Cells Identifier (PCI) per l'ottimizzazione delle reti mobili 4.5G e 5G. Inoltre, dato che il modello riguarda un caso reale e quindi complesso, risulta fondamentale utilizzare il Quantum Annealer applicando tutte le possibili strategie che ne consentono il miglior utilizzo. Questo si rende necessario in quanto modalità diverse impattano significativamente sulle caratteristiche della soluzione trovata, come approfondito in questo lavoro in pubblicazione su Springer Quantum Machine Intelligence [8].

Nel caso della pianificazione dei PCI, la velocità di esecuzione dell'algoritmo unito alla possibilità di trovare soluzioni di qualità molto alta a problemi non-lineari ha permesso di verificare come gli algoritmi Quantistici possano essere applicati con successo a problemi reali.

L. Asproni l.asproni@reply.it
D. Caputo da.caputo@reply.it
M. Magagnini m.magagnini@reply.it

l'assegnazione rispettando i vincoli di "riuso".

È quindi un problema di natura combinatoria analogo, in linea di principio, al problema di colorazione delle mappe geografiche, nelle quali due paesi confinanti devono assumere colori diversi.

Nella classe dei problemi di ottimizzazione combinatoria, un framework che trova ampia applicazione è il QUBO (Quadratic Unconstrained Binary Optimization)[7].

In questi problemi, generalmente, intervengono una serie di variabili binarie, che possono quindi assumere due soli valori corrispondenti a due stati "netti" (0/1, si/no, on/off...). Ogni combinazione del set di variabili determina un risultato che, a seconda del problema, rappresenta un costo o un guadagno.

In termini generali, il QUBO è un modello matematico per ottimizzare un problema che può essere rappresentato nella seguente forma:

minimizzare $y = x^T Q x$ [Eq 3]

Analizzando la formulazione matematica si può notare che:

- il problema viene descritto matematicamente in una forma quadratica, nella quale il vettore x rappresenta il set delle variabili di cui occorre trovare la combinazione di 0 e 1 ottimale
- la matrice Q è composta da elementi di valore costante che, oltre a modellare matematicamente il problema, tengono

conto anche dei vincoli a cui le variabili devono sottostare. L'artificio matematico per includere questi vincoli porta ad avere ad una formulazione "compatta", autoconsistente e quindi "unconstrained"

- l'ottimizzazione consiste nell'individuare il minimo della funzione di costo quadratica.

Mettendo assieme gli elementi descritti, si ottiene il "Quadratic Unconstrained Binary Optimization" model, cioè il QUBO.

La generalizzazione del modello ai casi in cui anziché il minimo occorra individuare il massimo, si deriva riformulando la funzione obiettivo del problema nel suo opposto negativo; in questo modo il metodo di risoluzione continua ad essere lo stesso, perché il valore minimo che si ottiene corrisponde al valore massimo del problema originario.

L'attività di sviluppo di un algoritmo di ottimizzazione si può così ricondurre alla riformulazione del problema secondo il modello QUBO e all'utilizzo delle librerie che risolvono questi casi in maniera efficiente. Questa metodologia di lavoro consente di standardizzare e rendere efficiente l'intero processo.

La grande varietà di problemi a cui il modello è applicabile, unitamente all'ampia gamma di piattaforme (CPU e GPU) che supportano le librerie software di risoluzione, danno al framework QUBO un ruolo distintivo

nell'ambito dei problemi di ottimizzazione combinatoria.

A rafforzare ulteriormente l'importanza del modello, contribuisce il fatto che anche nelle librerie software delle piattaforme di quantum computing sono stati integrati solver di risoluzione per il QUBO.

Per quanto visto nell'introduzione, i quantum annealer sono particolarmente adatti per i problemi di ottimizzazione di natura combinatoria; per questo motivo l'algoritmo di pianificazione PCI, sviluppato secondo il modello QUBO, è stato eseguito sul quantum computer di D-Wave 2000QTM.

Lo stesso algoritmo QUBO è stato elaborato anche su CPU e GPU.

Modellizzazione pianificazione identificativi di celle

Riprendendo la formulazione matematica degli identificativi di cella [Eq 1] unitamente ai vincoli di "riuso", la modellizzazione del problema – secondo il framework QUBO – si traduce nella seguente forma:

$$QUBO = \lambda_1 H_{PCI_{HC}} + \lambda_2 H_{GROUP_{HC}} + \lambda_3 H_{PCI_{SC}} + \lambda_4 H_{GROUP_{SC}} + \lambda_5 H_{GROUP_{YHC}} \quad [Eq 4]$$

Nella quale intervengono 5 contributi:

$H_{PCI_{HC}}$ esprime il vincolo di utilizzare PCI diversi per celle adiacenti

$H_{GROUP_{HC}}$ esprime il vincolo di utilizzare Group_ID differenti per nodi diversi

$H_{PCI_{SC}}$ esprime il costo che deriva dall'utilizzo dello stesso PCI su celle adiacenti

$H_{GROUP_{SC}}$ esprime il costo che deriva dall'utilizzo dello stesso Group_ID su nodi adiacenti

$H_{GROUP_{YHC}}$ esprime il vincolo di utilizzare Cell_ID diversi per celle dello stesso nodo.

I pesi di ognuno (le λ) devono essere calibrati per individuare l'area ottimale di lavoro.

L'elaborazione di questo algoritmo "completo", comprensivo di tutte le componenti, ha però presentato alcuni problemi:

- la calibrazione dei pesi (le λ) è risultata complessa per la quantità di parametri (5)
- il modello è poco scalabile
- la qualità della soluzione non è migliore rispetto al metodo tradizionale.

Queste criticità hanno portato così allo sviluppo di un nuovo modello seguendo una strategia alternativa:

- il modello è stato scomposto in fasi computazionali, definite in modo che la complessità computazionale di ognuna fosse gestibile con i computer classici e quantistici attuali, utilizzandoli anche entrambi con flussi di elaborazione ibrida
- le λ sono state suddivise nelle varie fasi, semplificandone il tuning.

Così facendo l'algoritmo QUBO è stato strutturato in due fasi:

1. primo layer: assegnazione degli identificativi di gruppo (Group_ID) con la definizione di un modello QUBO che coinvolge i soli due termini: $H_{GROUP_{HC}}$, $H_{GROUP_{SC}}$
2. secondo layer: assegnazione dei Cell_ID di ogni cella all'interno dei gruppi, utilizzando l'assegnazione dei gruppi del primo layer, ottenendo così la distribuzione dei PCI di ogni cella. Il corrispondente modello QUBO contiene quindi i tre termini: $H_{PCI_{HC}}$, $H_{GROUP_{YHC}}$, $H_{PCI_{SC}}$

I due layer così costituiti, hanno singolarmente una complessità computazionale inferiore rispetto all'algoritmo "completo".

Con questo nuovo modello è stato possibile elaborare set di celle di dimensioni significative, ottenendo una qualità migliore rispetto alla procedura di pianificazione dei PCI basata su tecniche combinatorie tradizionali.

Una sintesi dei risultati ottenuti è riportata nella sezione sui risultati dell'elaborazione.

Segmentazione dei dati di input

Lo sviluppo dell'algoritmo non si è limitato alla sola modellizzazione QUBO, ma è stato necessario tarare i set di celle affinché le loro dimensioni fossero gestibili con le potenze di

calcolo dei computer a disposizione (CPU, GPU, QPU).

Questo perché la complessità computazionale del problema cresce esponenzialmente all'aumentare del numero delle celle.

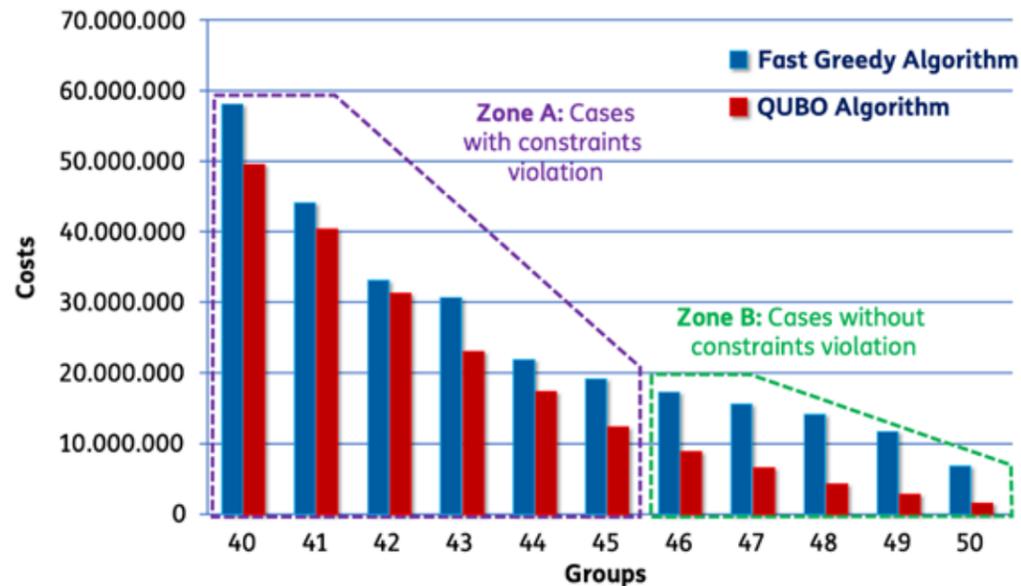
In funzione dei vincoli sulla potenza di calcolo, sono stati creati dei set di dimensioni massime di alcune centinaia di celle, sulla base delle loro adiacenze.

Risultati dell'elaborazione dell'algoritmo QUBO di pianificazione PCI

L'applicazione dell'algoritmo QUBO al contesto della pianificazione dei PCI delle reti 5G e LTE ha fornito risultati incoraggianti dal confronto con quelli garantiti dal Fast Greedy Algorithm già utilizzato nell'ambito del framework Open SON di TIM. Dal punto di vista dei tempi di calcolo necessari per definire un piano dei PCI, non è facile effettuare un confronto "omogeneo" poiché i criteri di stop dei due algoritmi sono profondamente diversi. Ciononostante, da misure indicative e tenendo conto del processo complessivo emerge la chiara indicazione di una riduzione del tempo di un fattore 10x, con la prospettiva di ulteriori margini di miglioramento.

La riduzione dei tempi di calcolo è di particolare interesse nell'ottica dell'evoluzione della pianificazione

PCI planning costs comparison



4
Grafico di confronto su insieme campione

dei PCI verso modalità “real-time” (ad esempio attraverso l’integrazione in sistemi di tipo RIC – Radio Intelligent Controller nell’ambito delle architetture definite dalla O-RAN Alliance).

Oltre ai tempi di calcolo, al fine di analizzare le potenzialità dell’algoritmo su piani di complessità variabile, sono state effettuate una serie di prove diminuendo il numero di gruppi disponibili (rispetto al massimo consentito dallo standard).

Al diminuire del numero di gruppi, infatti, aumenta la probabilità di violazione dei vincoli. In tal modo, è stato individuato il minimo numero di PCI necessario per definire un piano ottimale, cioè tale da soddisfare tutti i vincoli di importanza “prioritaria” (relativi al PCI completo) e sono stati

ottenuti i costi complessivi del piano (sia “primari”, sui PCI, sia “secondari”, sul GROUP_ID, valutati mediante la [Eq 2]) quando questo viola uno o più vincoli di “riuso” sull’identificativo di cella completo.

A titolo di esempio, si presentano i risultati ottenuti nel caso di un insieme campione di 450 celle della rete TIM, riportati in Figura 4:

Come si può osservare, in tutti i casi si ha un miglior comportamento dell’algoritmo QUBO.

Nella zona A, quest’ultimo ottiene soluzioni con un numero di vincoli inferiore rispetto al precedente algoritmo Fast Greedy (che conferma comunque la propria efficacia), mentre nella zona B – cioè in assenza di

vincoli violati – esso garantisce funzioni costo inferiori grazie al minor numero di Group_ID utilizzati (46 contro 50) e ad una più efficace gestione del vincolo di “riuso” relativo al GROUP_ID.

Verso un’applicazione sistematica del Quantum Computing

Allo stato attuale, sono in corso sviluppi dell’algoritmo di pianificazione dei PCI mirati a far evolvere l’algoritmo QUBO in modo da gestire insieme sempre maggiori di celle.

Con questo obiettivo è stata definita una pipeline di lavoro, articolata su tre fasi:

- una di clustering da realizzare attraverso una procedura di decomposizione dei set di celle in base alle loro adiacenze
- due fasi successive che corrispondono ai due livelli di elaborazione presentati nella sezione descrittiva del modello prevedendo anche il tipo di hardware (CPU, GPU, QPU) più appropriato per ogni fase, rispetto alla complessità computazionale che dovranno gestire.

Nel secondo layer di elaborazione è prevista, in particolare, l’invocazione al quantum computer di D-Wave.

L’architettura della pipeline è schematizzata nella figura 5.

La pianificazione dei PCI della rete mobile è stato il primo caso d’uso affrontato in TIM con un approccio

quantistico, utile per verificare l’applicabilità, le implicazioni e i vantaggi di questa nuova metodologia di programmazione.

È importante evidenziare come in ambito “Telco” sia possibile individuare ulteriori problemi adatti a tale tipologia di approccio e come altri use case siano stati identificati come potenziali candidati per la modellizzazione secondo la metodologia quantum.

Se gli studi di fattibilità daranno esito positivo – un possibile esempio sono gli algoritmi di Coverage and Capacity Optimization, anche ove basati sull’impiego dell’AI – questi casi d’uso verranno trasportati nel mondo della computazione quantistica.

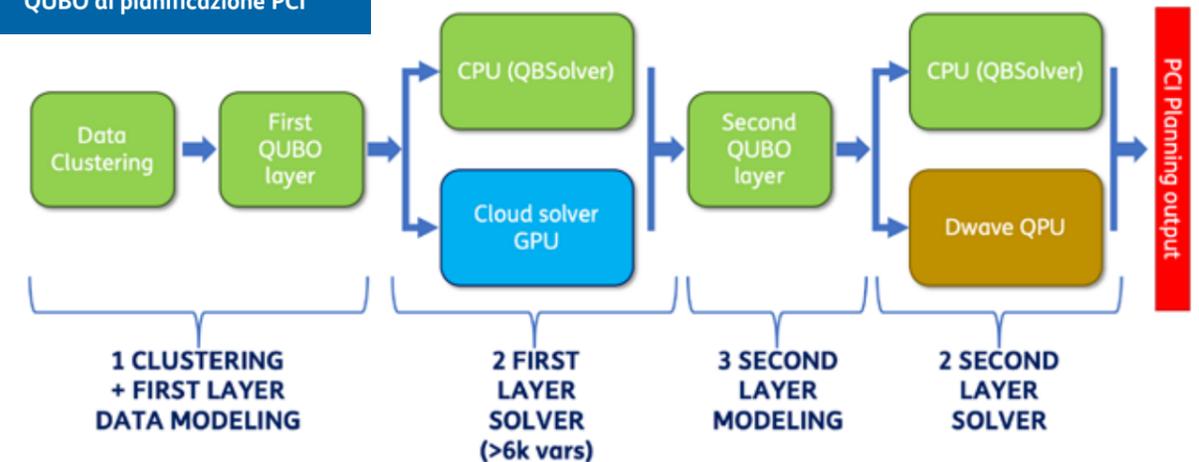
Tra gli use case candidati si stanno considerando alcune applicazioni non real-time e near real-time pro-

prie dei RAN Intelligent Controller (RIC) che, secondo il paradigma della open RAN, potrebbero integrarsi con le funzionalità real-time proprietarie dei vendor, con le quali opererebbero in sinergia a fini di ottimizzazione della gestione delle risorse radio.

Un fattore che favorirà l’utilizzo del quantum è legato al possibile impiego on-premises, per gestire la complessità delle applicazioni time-critical e tra queste, come già accennato, il suo utilizzo a livello di edge.

L’obiettivo di un’azienda come TIM, è – in conclusione – quello di estendere ulteriormente l’utilizzo del quantum computing alle applicazioni in ambito Telco, per gestire e risolvere problemi reali, tranne beneficio e migliorare la qualità dei servizi offerti alla clientela grazie a questa tecnologia innovativa.

5
Pipeline obiettivo di evoluzione dell’algoritmo QUBO di pianificazione PCI



Bibliografia

1. State of Quantum Computing in 2020 for Business Leaders. (s.d.). Tratto da Blog Aimultiple: <https://blog.aimultiple.com/quantum-computing/>
2. [2] KATWALA, A. (s.d.). Quantum computers will change the world (if they work). Tratto da Wired: <https://www.wired.co.uk/article/quantum-computing-explained>
3. [3] Quantum Computing Applications in 2020. (s.d.). Tratto da Blog AIMultiple: <https://blog.aimultiple.com/quantum-computing-applications/>
4. [4] Bertotto, P., Epifani, F., Ludovico, M., & Zarba, G. (s.d.). Open Self-Organizing Network: a continuous development for Radio Access Network performance optimization. Tratto da Telecom Italia: <https://www.telecomitalia.com/tit/it/notiziariotecnico/edizioni-2019/n-3-2019/N3-Open-Self-Organizing-Network-continuous-development-for-Radio-Access-Network-performance-optimization.html>
5. [5] Bertotto, P., & Zarba, G. (s.d.). Applicazioni "Open SON" nella rete TIM. Tratto da <https://www.telecomitalia.com/content/tiportal/it/notiziariotecnico/edizioni-2018/n-1-2018/N6-DigiRAN-valore-automazione-accesso-radio/appfondimenti-2.html>
6. [6] B., C. (s.d.). The Greedy Algorithms Class: Formalization, Synthesis and Generalization. Tratto da Université catholique de Louvain: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.34.5676&rep=rep1&type=pdf>
7. [7] Glover, F. K. (2019, May). A Tutorial on Formulating and Using QUBO Models. Tratto da Leeds Faculty of Colorado: <http://leeds-faculty.colorado.edu/glover/511%20-%20QUBO%20Tutorial%20-%20updated%20version%20-%20May%202019.pdf>
8. [8] Asproni, et al., Accuracy and minor embedding in subqubo decomposition with fully connected large problems: a case study about the number partitioning problem, arXiv:1907.01892v2



Andrea Boella andrea.boella@telecomitalia.it

Ingegnere elettronico, è in Telecom Italia dal 1997, dove è entrato nella ex direzione territoriale del Piemonte e Valle d'Aosta iniziando con la pianificazione della rete di trasporto regionale, per passare in qualità di responsabile prima al supporto specialistico trasmissioni e successivamente al secondo livello di assistenza tecnica per la clientela enterprise. Dal 2003 al 2005, come espatriato, è stato responsabile del settore esercizio in Telecom Italia France e al ritorno in Italia entra a far parte del gruppo ex Global Network coordinando, come project manager, attività di supporto verso le partecipate estere inizialmente su tematiche di rete fissa e a partire dal 2008 sulla core network di rete mobile. Da fine 2018 lavora nel gruppo di Service Innovation con attività di ricerca su servizi per l'IOT, blockchain, quantum computing e quantum communication ■



Michele Ludovico michele.ludovico@telecomitalia.it

Ingegnere elettronico, ha iniziato ad operare nel gruppo TIM progettando sistemi a micro-onde per comunicazioni via satellite. Dal 2001 si occupa di strumenti e metodologie di progettazione ed ottimizzazione dell'accesso radio, che TIM sviluppa "in house" a supporto dell'evoluzione della rete mobile. Dal 2014 è responsabile della funzione di TIM che assicura l'ingegneria e lo sviluppo delle soluzioni di automazione per la rete di accesso radio, secondo il paradigma "Self Organizing Network". Ha svolto, inoltre, attività di formazione e consulenza in Italia ed all'estero ed è co-inventore di diversi brevetti nel campo della progettazione wireless e della gestione delle risorse radio ■



Giuseppe Minerva giuseppe.minerva@telecomitalia.it

Laurea in Ingegneria Elettronica nel 1991 e assunzione nel gruppo TIM nel marzo del 1992. Servizio prestato nei settori di Innovazione e di Ingegneria di Rete. Attività nel campo della definizione di algoritmi e metodologie di pianificazione e ottimizzazione dell'accesso radio fin dai tempi della nascita della rete GSM e dell'ideazione dell'algoritmo di assegnazione frequenziale usato da TIM. Attività di formazione interna e di brevettazione nel campo della progettazione radio. Attualmente coinvolto nella transizione delle metodologie di pianificazione e ottimizzazione della rete TIM verso i paradigmi SON e di Full-Automation ■



Mauro Alberto Rossotto mauro.rossotto@telecomitalia.it

Laureato all'Università di Torino in Informatica con specializzazione Intelligenza Artificiale. Entrato in Telecom Italia nel 1995 ha partecipato a diversi progetti realizzativi legati a Data Mining Lab, analisi dati a scopo Antifrode e Marketing, Push-to-talk, Smart Inclusion. Dal 2012, in Innovation, ha seguito in qualità di responsabile di struttura le attività legate allo sviluppo di servizi innovativi su device connessi, a partire da TIM Vision e TIM Cloud su Smart TV e Game Console.

Nel 2014 in Strategy & Innovation sono iniziate le prime attività sul mondo IoT, ed in particolare sui verticali Smart Home, Wellness, Smart City, Smart Retail, Energy e Industry affrontando tutti gli aspetti tecnologici del servizio. Oggi in Service Innovation è responsabile del Program "Internet of Everything" dove vengono seguite tematiche relative a Droni, Robotics, IoT, Blockchain e Quantum Technologies con attività di scouting, prototipazione, sperimentazione e pipeline verso ingegneria ■