Agenzia nazionale per le nuove tecnologie,
l'energia e lo sviluppo economico sostenibile

MINISTERO DELLO SVILUPPO ECONOMICO

Ricerca di Sistema elettrico

Analisi della gestione dei dati, protocolli di trasmissione e vulnerabilità negli scenari di smart street

F. Riganti Fulginei, S. Panzieri, R. Torlone, D. Firmani, F. Pascucci, G. Bernieri

ROMA TRE
UNIVERSITÀ DEGLI STUDI

Report RdS/PAR2015/009

ANALISI DELLA GESTIONE DEI DATI, PROTOCOLLI DI TRASMISSIONE E VULNERABILITÀ NEGLI SCENARI DI SMART STREET

F. Riganti Fulginei, S. Panzieri, R. Torlone, D. Firmani, F. Pascucci, G. Bernieri (Università Roma Tre)

Settembre 2016

Report Ricerca di Sistema Elettrico

Accordo di Programma Ministero dello Sviluppo Economico - ENEA
Piano Annuale di Realizzazione 2015
Area: "Efficienza energetica e risparmio di energia negli usi finali elettrici e interazione con altri vettori energetici"
Progetto: D5 "Innovazione tecnologica, funzionale e gestionale nella illuminazione pubblica ed in ambienti confinati"
Obiettivo: B "Smart Street"
Responsabile del Progetto: Nicoletta Gozo, ENEA

Il presente documento descrive le attività di ricerca svolte all'interno dell'Accordo di collaborazione "Analisi della gestione dei dati, protocolli di trasmissione e vulnerabilità negli scenari di *smart street*"
Responsabile scientifico ENEA: Stefano Pizzuti
Responsabile scientifico Università Roma Tre: prof. R. Torlone

Indice

SOMMARIO 4
1 INTRODUZIONE 5
2 DESCRIZIONE DELLE ATTIVITÀ SVOLTE E RISULTATI 5
2.1 BIG DATA PER GLI SMART STREET 5
2.1.1 Tecnologie per la gestione dei big data 5
2.1.2 Una architettura big data per lo smart street 15
2.1.3 Tipologie di analisi 17
2.1.4 Analisi batch 18
2.1.5 Analisi real-time 19
2.1.6 Data cleaning 20
2.1.7 Migrazione nel nuovo sistema 20
2.2 CYBER SECURITY IN UNO SCENARIO SMART STREET 21
2.2.1 Inquadramento 21
2.2.2 Power Line Communication 22
2.2.3 Comunicazione a radiofrequenza 29
2.2.4 Vulnerabilità dei sistemi PLC 29
2.2.5 Criticità dei sistemi di monitoraggio e controllo utilizzando internet 34
2.2.6 Il Sistema Smart Street in ENEA 39
2.2.7 Architettura del sistema di controllo remoto e dispositivi 40
2.2.8 Topologia della rete 43
2.2.9 Sicurezza del sistema Smart Street 45
2.3 I PROTOCOLLI DI COMUNICAZIONE DEGLI SMART STREET 47
2.3.1 Introduzione al protocollo TALQ 47
2.3.2 Messaggi - Descrizione 50
2.3.3 Messaggi - Supporto 51
2.3.4 Modello Dati 53
2.3.5 Funzioni 54
2.3.6 Dispositivi Logici 60
2.3.7 Servizi 61
2.3.8 Eventi 62
2.3.9 Sincronizzazione Dati 62
2.3.10 Gestione delle entità 63
2.3.11 Sincronizzazione dispositivi logici 65
2.3.12 Scenario Applicativo – Smart Street con Protocollo TALQ 66
2.3.13 Studio Protocollo TALQ - Conclusioni e sviluppi 67
3 CONCLUSIONI 68
4 RIFERIMENTI BIBLIOGRAFICI 68
5 CURRICULUM SCIENTIFICO DEL GRUPPO DI LAVORO 70

Sommario

In questo documento sono affrontate tre importanti problematiche relative alla innovazione tecnologica, funzionale e gestionale della illuminazione pubblica (smart lithing) e degli ambienti confinati (smart street):

- **i big data**, ovvero la definizione dei meccanismi di gestione e analisi di grandi quantità di dati di carattere energetico provenienti dai sistemi di illuminazione pubblica,

- **la cyber security**, ovvero lo studio delle possibili vulnerabilità dei sistemi di illuminazione pubblica in ambito urbano,

- **i protocolli di comunicazione**, ovvero l'analisi dei meccanismi standard per la trasmissione di informazioni riguardanti l'illuminazione pubblica.

Per ciascuno di questi aspetti viene fatto un inquadramento generale del problema, vengono affrontati i principali aspetti metodologici e tecnologici e vengono proposte delle soluzioni nello specifico scenario della Smart Street installata presso la sede di ENEA – Casaccia.

1 Introduzione

Il lavoro presentato in questo documento si inquadra nell'ambito dell'accordo di programma tra il Ministero dello Sviluppo Economico e l'ENEA che riguarda la tematica generale dell'efficienza energetica e risparmio di energia negli usi finali elettrici e nell'interazione con altri vettori energetici. In questo ambito, lo specifico progetto di interesse riguarda "l'innovazione tecnologica, funzionale e gestionale nella illuminazione pubblica ed in ambienti confinati" e l'obiettivo è quello della definizione di metodi e strumenti per la realizzazione degli "Smart Street", anche detti "Smart Lighting".

Con il termine *Smart Lighting* (SL) si indicano tutte quelle tecnologie di illuminazione progettate per essere efficienti dal punto di vista elettronico. Questo comprende sia impianti ad alta efficienza che impianti automatizzati in modo da adattare le condizioni di illuminazione alla presenza di utenti o in base alle condizioni di luminosità ambientale.

Nei sistemi di SL legati all'illuminazione pubblica, così come in ogni declinazione delle applicazioni di Smart Cities (SC), le comunicazioni giocano un ruolo essenziale, essendo tali sistemi distribuiti, complessi e con numerosi punti di accesso. Le informazioni relative ai sensori, ai contatori intelligenti, agli attuatori devono arrivare ai nodi di controllo dove vengono implementate le strategie che rendono più vivibili le nostre città.

Come si è potuto notare, i sistemi di comunicazione rappresentano un anello fondamentale nella realizzazione di un sistema SL. Questi sistemi, per la loro natura, sono anche i sistemi più soggetti ad attacchi. Il presente documento si pone l'obiettivo di analizzare i sistemi di SL con particolare riferimento alla Smart Street installata presso la sede di ENEA – Casaccia.

L'obiettivo generale è mostrare come le soluzioni proposte possono contribuire a migliorare l'efficacia e l'efficienza energetica e il risparmio di energia negli usi finali.

2 Descrizione delle attività svolte e risultati

2.1 Big data per gli smart street

L'obiettivo finale di questo task consiste nel progetto e nella sperimentazione di una piattaforma informatica flessibile e scalabile in grado di raccogliere, gestire e analizzare dati di grandi dimensioni provenienti da diverse tipologie di sistemi sensoriali installati presso l'ENEA. Lo scenario di riferimento sarà quello dei dati provenienti da una rete di illuminazione pubblica di strade ma si intende delineare un approccio che abbia validità generale e che sia quindi applicabile anche alla raccolta e l'analisi di dati provenienti da sensori installati in contesti diversi, per esempio in edifici o in altre infrastrutture.

Il sistema deve garantire il supporto alla diagnostica avanzata degli smart street e in particolare alla definizione di strategie di controllo e ottimizzazione della illuminazione pubblica. L'applicazione è inquadrabile nell'ambito dei cosiddetti "big data per l'IoT" perché i dati arrivano dai sensori in flussi, sono raccolti in maniera incrementale, possono raggiungere dimensioni considerevoli e richiedono attività di monitoraggio e analisi che coinvolgono grandi porzioni dei dati memorizzati.

2.1.1 Tecnologie per la gestione dei big data

Si parla di "big data" quando i dati da gestire superano le capacità di memorizzazione, gestione e analisi tipiche dei tradizionali sistemi per le basi di dati. Le caratteristiche generali di questi enormi magazzini di dati vengono spesso descritte secondo tre dimensioni, dette le tre "v" dei big data, come viene illustrato in Figura 1.

- Volume: dimensione dei dati da gestire misurata non solo in termini assoluti, ma anche in termini di andamento di crescita e di requisiti di prestazioni per la loro elaborazione.

- Varietà: tipologia dei dati e delle sorgenti; nella maggior parte dei casi si tratta di dati semi-strutturati (per esempio log memorizzati su file in formato XML, Json, CVS o in formati proprietari) o destrutturati (*raw text* che possono essere pagine Web, file di testo, documenti) per i quali tipicamente non esiste uno schema.

- Velocità: rapidità con la quale i dati arrivano e devono essere elaborati; alcune applicazioni consentono la elaborazione batch, ma in molti casi bisogna operare il real-time e non sono rari casi in cui i dati viaggiano in stream, ovvero in flussi, che vanno elaborati alla velocità nella quale arrivano.

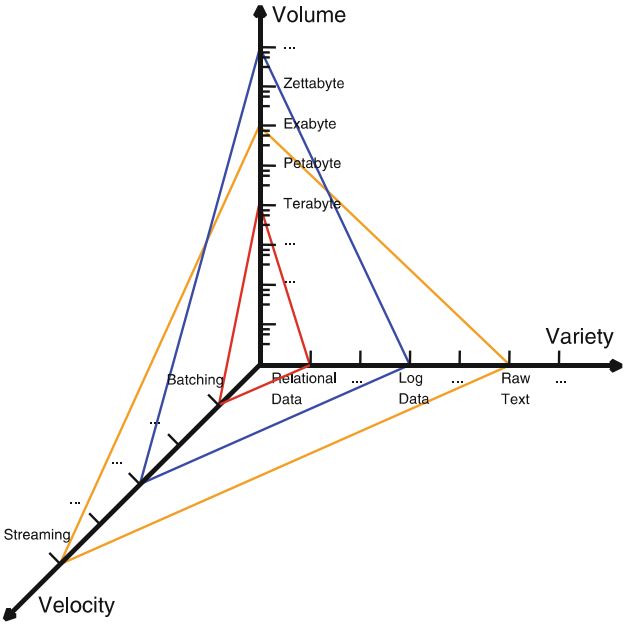


Figura 1 – Le dimensioni dei big data.

A causa di queste caratteristiche, nell'ambito della gestione di big data si stanno recentemente diffondendo delle soluzioni tecnologiche alternative a quelle tradizionali che cercano di soddisfare i seguenti requisiti:

- si adattano meglio a contesti applicativi come quello delle smart street,

- garantiscono prestazioni migliori,

- consentono la scalabilità delle applicazioni su grandi moli di dati che crescono rapidamente,

- sono in grado di bilanciare latenza (tempi di risposta di una query o di una analisi), throughput (numero di operazioni svolte in un intervallo di tempo) e tolleranza ai guasti (affidabilità in presenza di malfunzionamenti software e hardware).

È difficile a tale riguardo definire architetture precise a causa delle innumerevoli soluzioni che sono state proposte e della grande varietà dei domini applicativi ognuno dei quali richiede un approccio diversificato. È però possibile individuare alcune soluzioni di carattere generale che, anche se non costituiscono soluzioni

imprescindibili, sono molto comuni nelle applicazioni reali per big data e sono particolarmente adatte al settore IoT.

- I dati vengono distribuiti e replicati su cluster di computer per garantire la scalabilità delle applicazioni e per aumentare la tolleranza ai guasti e la disponibilità dei dati. La tecnologia di base che si è ormai affermata come standard di riferimento è quella di Apache Hadoop, un framework open-source per l'elaborazione distribuita basato sul file-system HDFS.

- Le operazioni richieste dall'analisi dei dati vengono seguite in parallelo in ambienti di elaborazione distribuita. Si può usare allo scopo il paradigma di programmazione MapReduce, che è nativo in Hadoop e che consente l'esecuzione di algoritmi paralleli su cluster di computer. Recentemente, si stanno affermando framework per l'elaborazione distribuita più efficienti ed evoluti rispetto a MapReduce, come Apache Spark e le soluzioni "SQL over Hadoop".

- I dati vengono memorizzati in formati semi o de-strutturati di cui non è in genere disponibile lo schema. In molti casi si fa uso per questo di un semplice file system distribuito, come HDFS. In alternativa, si usano sistemi NoSQL che poggiano su HDFS e offrono alle applicazioni delle librerie di più alto livello per l'accesso ai dati (per esempio per gestire la concorrenza) sulla base formati elementari, per esempio mediante semplici coppie chiave–valore.

- Le risorse hardware e software vengono virtualizzate adottando il paradigma del cloud-computing. Sul mercato esiste un'ampia disponibilita di offerte in questo ambito provenienti da aziende leader nel settore, come Amazon e Microsoft. I modelli di servizio offerti possono essere diversi, andando dall'Infrastructure as a Service (IaaS), nella quale solo l'infrastruttura è virtualizzata, al Software as a Service (SaaS), nel quale si virtualizza anche il software, che risiede e opera sul cloud.

Nel seguito di questa sezione verranno brevemente illustrate in maggior dettaglio le principali tecnologie appena citate. Questa panoramica consentirà di comprendere, con maggior chiarezza, le scelte architetturali proposte per lo scenario in esame. Una premessa importante da fare è che tutti gli strumenti software citati nel seguito sono di tipo open source e non richiedono pertanto costi di licenza. La maggior parte di essi appartengono al progetto della Apache Software Foundation (https://www.apache.org/).

Hadoop e MapReduce

Apache Hadoop (http://hadoop.apache.org/) è un progetto open-source della Apache Software Foundation per memorizzazione e la elaborazione di grandi moli massive di dati in ambienti distribuiti basati su cluster di computer. Il principio ispiratore di Hadoop è quello della "scalatura orizzontale" che prevede di aumentare lo spazio di memorizzazione e la potenza elaborativa di una infrastruttura hardware aggiungendo semplicemente a una rete di computer (comuni workstation dotate di risorse hardware standard) nuovi elementi sui quali distribuire il carico di lavoro, piuttosto che aumentare continuamente la capacità elaborativa di un unico sistema centralizzato ("scalatura verticale") come illustrato in Figura 1a.

Oltre al vantaggio di sapersi adattare rapidamente a dati di dimensione via via crescente, questa soluzione offre anche una ottima resistenza ai guasti perché i dati vengono replicati più volte su nodi diversi del cluster e quindi il malfunzionamento di un nodo non comporta né perdita di dati né interruzioni o rallentamenti di elaborazioni in corso.

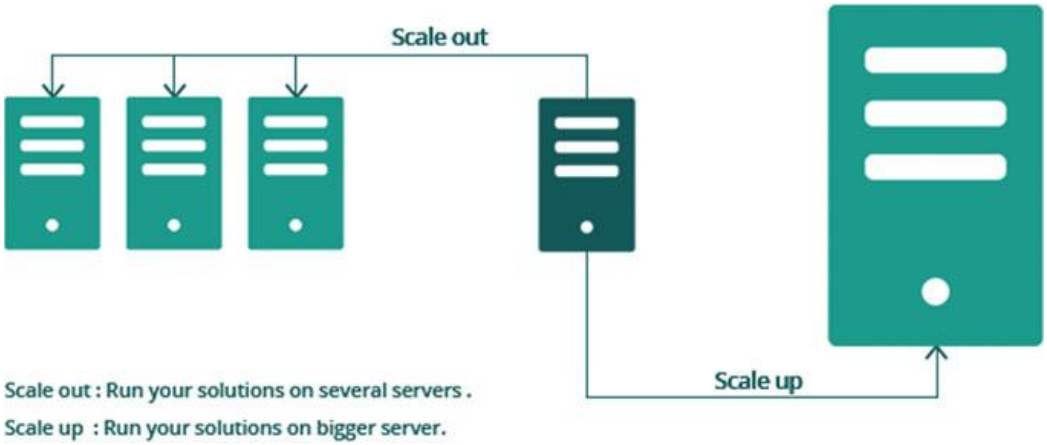


Figura 1a – Scalatura orizzontale e verticale a confronto

Hadoop consiste di due componenti principali:

- Hadoop Distributed File Sytem (HDFS): è un file system distribuito scritto in Java ed eseguibile su un cluster di macchine. HDFS è in grado di memorizzare file di grandi dimensioni frammentandoli sui nodi del cluster e gestendo l'affidabilità mediante la replicazione dei dati. Un cluster Hadoop contiene un nodo principale (detto NameNode) e diversi nodi controllati (detti nodi worker). Il nodo principale gestisce la distribuzione di blocchi di un file sui diversi nodi controllati.

- Hadoop MapReduce: fornisce un modello di elaborazione parallela, ispirato all'omonimo progetto originalmente sviluppato da Google. Un processo MapReduce (job) è composta da: (i) una componente di Map, che applica a ciascun record del file in ingresso una operazione (per esempio una semplice selezione) e genera per ciascun record una coppia (chiave,valore), e (ii) una componente Reduce che applica a ciascun gruppo di record con la stessa chiave generato dalla Map una funzione aggregativa (per esempio una somma). L'esecuzione è governata da un processo, detto job tracker allocato sul NameNode del cluster: il job tracker decompone un job in unità di lavoro (task) e assegna i task ai nodi sui quali il file di input è distribuito (secondo il principio di muovere la elaborazione sui dati invece che viceversa), controlla la terminazione dei task e riassegna quelli che falliscono.

Figura 2 illustra un semplice esempio di elaborazione MapReduce che calcola la frequenza di ciascuna parola in un file di testo. Il file è suddiviso in blocchi che vengono distribuiti e replicati su HDFS. La Map si occupa di generare per ciascuna parola w nel blocco su cui opera la coppia (w,1). In una fase intermedia vengono creati dei gruppi per ciascuna parola. Infine la Reduce somma i valori numerici di ciascun gruppo restituendo la parola e il risultato della somma.

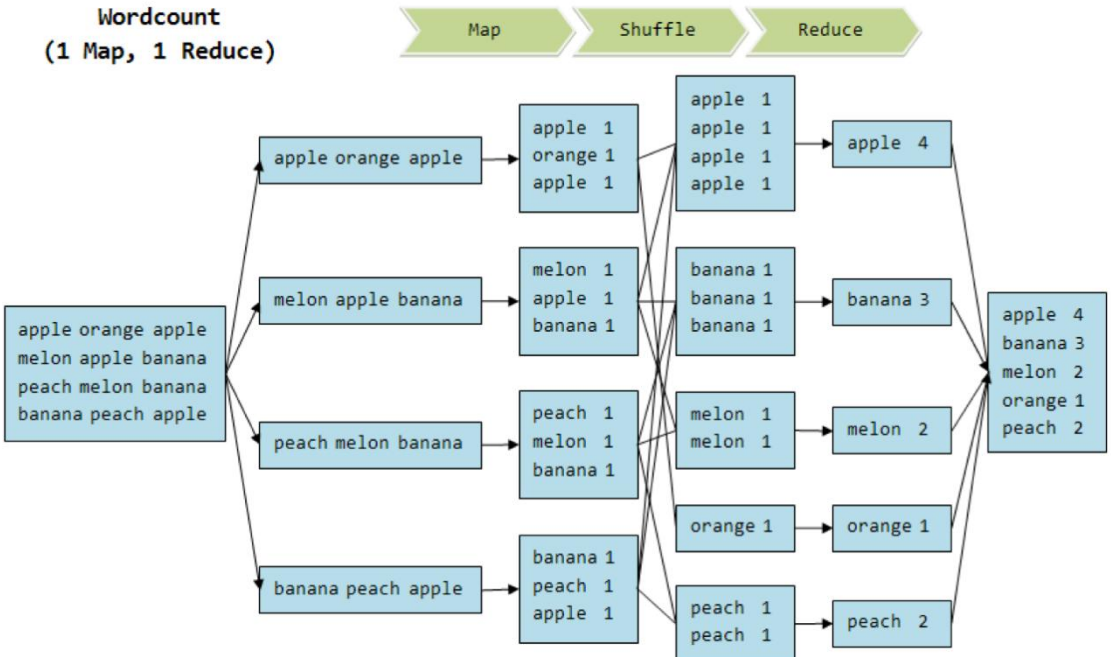


Figura 2 – Conteggio di parole in un file con MapReduce

Spark

Il paradigma MapReduce è particolarmente efficace per elaborare analisi tipiche sui big data ed ha avuto per questo molto successo. Presenta tuttavia alcune criticità che ne limitano l'efficienza, in particolare la necessità di leggere e scrivere i risultati intermedi di un job su memoria secondaria.

Apache Spark (http://spark.apache.org/) è un altro progetto open-source per la elaborazione analitica di big data che sostanzialmente cerca di ovviare a questa limitazione di MapReduce sfruttando in maniera più efficace le memorie principali dei nodi di un cluster utilizzandole per la memorizzazione dei risultati intermedi, come illustrato in Figura 3. Inoltre, offre una serie di operazioni predefinite più ricco rispetto a MapReduce (oltre alla Map e alla Reduce esistono operazioni di filter, union, collect, ecc.) che rendono più flessibile la programmazione. Spark opera su dati distribuiti che possono essere memorizzati su HDFS (ma non necessariamente). Infine, oltre alle funzionalità citate che operano tipicamente in modalità batch, Spark offre anche uno specifico strumento, chiamato Spark streaming, che è in grado di effettuare analisi real-time su flussi continui di dati (stream) operando su dati aggregati su intervalli temporali molto brevi detti di micro-batch.

Per tutti questi motivi, mentre la componente HDFS di Hadoop è ancora uno standard di riferimento per la memorizzazione di grandi moli di dati in ambiente distribuito, Spark è recentemente diventato il sistema più utilizzato per la elaborazione di questi dati, anche rispetto a MapReduce che è nativo in Hadoop.

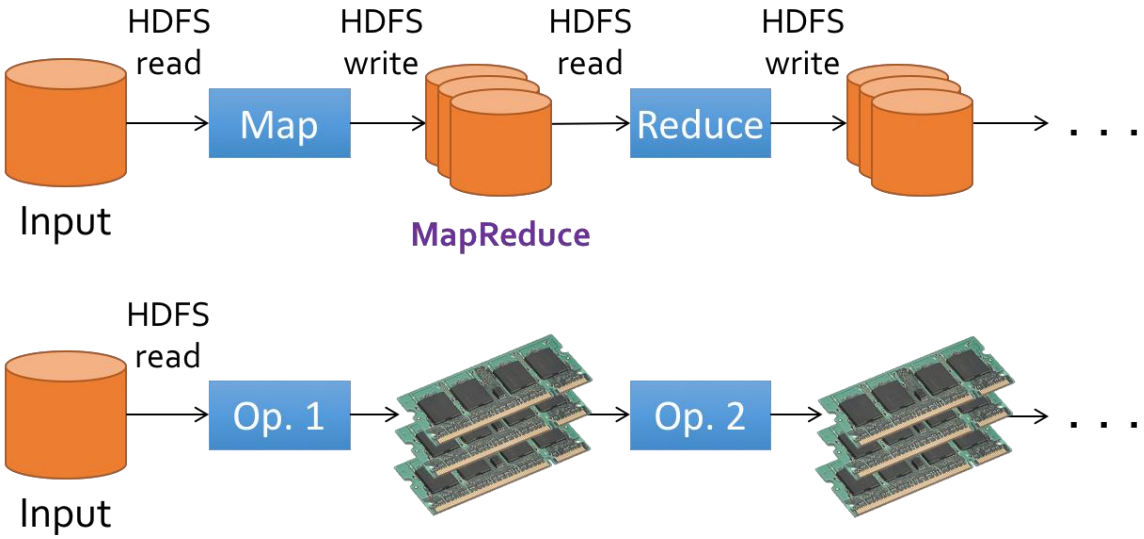


Figura 3 – Elaborazione MapReduce e Spark a confronto

SQL su Hadoop

Spark e MapReduce sono strumenti potenti per le analisi massive di big data ma non sono semplici da programmare e richiedono skill che sono difficili da trovare tra gli operatori del settore di data management. Per questo motivo si sono diffuse alcune tecnologie che offrono delle interfacce di più altro livello per la programmazione distribuita su cluster.

Apache Hive (https://hive.apache.org/) è uno degli strumenti più diffusi in questo contesto ed offre un'interfaccia SQL (uno storico linguaggio per l'interrogazione di database molto noto ai programmatori) per l'interrogazione e la manipolazione dei dati salvati su HDFS. In Hive, dati memorizzati su HDFS sono visti dal programmatore come tabelle di un database relazionale e un comando SQL standard su queste tabelle è tradotto automaticamente in una computazione MapReduce. Una serie di ottimizzazioni garantisce l'esecuzione efficiente del comando SQL.

Esistono diverse soluzioni analoghe ad Hive e per questo classificate generalmente come strumenti "SQL su Hadoop". Per esempio Pig, che usa una sintassi SQL-like, Spark SQL, che come suggerisce il nome si basa sul framework Spark invece che su MapReduce, e Impala, che poggia su HDFS ma usa un meccanismo di esecuzione di query SQL specializzato e autonomo. In effetti, l'idea di utilizzare la componente HDFS di Hadoop si è largamente diffusa: esistono infatti molti strumenti software in grado di offrire funzionalità diverse per la gestione e l'analisi di varie tipologie di big data che utilizzano HDFS per la distribuzione, la duplicazione, la memorizzazione e l'accesso ai dati (vedi Figura 4 nella quale YARN è l'interfaccia che Hadoop 2.0 espone su Web per permettere l'interazione con HDFS). Queste soluzioni vengono comunemente denominate "applicazioni su/per Hadoop" anche se il termine Hadoop è qui usato in maniera impropria, trattandosi in effetti della componente HDFS di Hadoop.

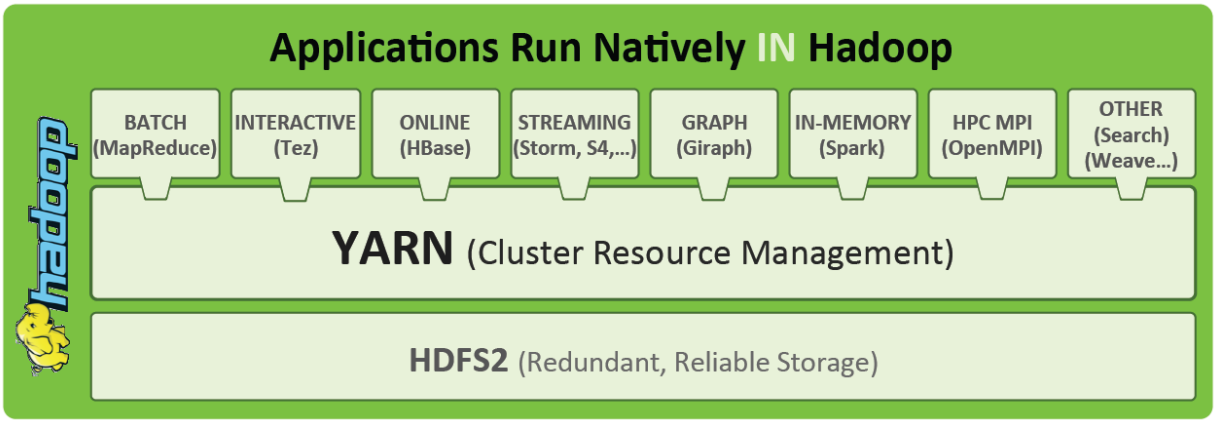


Figura 4 – Applicazioni che operano su HDFS

Sistemi NoSQL

Le soluzioni software sopra citate sono adatte alla elaborazione efficiente di big data ma non offrono funzionalità di persistenza e concorrenza tipiche dei sistemi classici per la gestione dei dati (i cosiddetti DBMS). Al tempo stesso i DBMS tradizionali non sono adatti a lavorare su cluster di computer ed è stato mostrato che non riescono a garantire prestazioni elevate quando la dimensione dei dati aumenta rapidamente fino a raggiungere dimensioni superiori a quelle tipiche dei tradizionali database. Inoltre, il modello relazionale sul quale i DBMS tradizionali si poggia non è adatto alle tipologie di dati tipiche degli scenari di big data, nei quali i dati sono tipicamente poco o per niente strutturati.

Per questi motivi si sono diffusi diversi sistemi per la gestione dei dati che utilizzano dei formati di memorizzazione dei dati non-relazionali e dei meccanismi per la loro manipolazione alternativi rispetto a quelli tipici dei DBMS relazionali. In particolare, questi sistemi non utilizzano il linguaggio standard dei sistemi relazionali e per questo motivo vengono comunemente chiamati sistemi NoSQL.

La caratteristica comune di questi sistemi è la capacità di gestire in maniera efficiente dati distribuiti su un cluster di computer e di garantire ottime prestazioni su grandi moli di dati al prezzo però di non riuscire a garantire, in maniera completa, tutte le funzionalità tipiche di un DBMS. In particolare, non garantiscono completamente le proprietà “acide” (Atomicità, Consistenza, Isolamento e Durabilità) delle transazioni. Questo limite è però relativamente poco rilevante in molti contesti applicativi dei big data (per esempio nel mondo IoT nel quale i dati registrati sono semplici misurazioni provenienti da sensori che non possiedono requisiti di gestione particolarmente stringenti). Per questo, questi sistemi hanno avuto un notevole successo negli ultimi anni.

Sebbene esistano decine di sistemi NoSQL molto diversi tra loro, è possibile definire una loro classificazione di massima:

- Document databases: i dati vengono manipolati sotto forma di un documento strutturato e nidificato, tipicamente utilizzando una sintassi standard chiamata Json. Le primitive di accesso ai dati consentono di interrogare il database sulla base di campi e sottocampi di questi documenti. Non esiste però uno schema di riferimento e neanche un linguaggio di manipolazione dati standardizzato per tutti i sistemi appartenenti a questa categoria. I document databases più popolari sono MongoDB e Couchbase.

- Graph databases: i dati vengono rappresentati in forma di grafo: tipicamente, i nodi del grafo descrivono oggetti e gli archi relazioni tra questi oggetti. Le operazioni di accesso ai dati si basano

su primitive di navigazione nel grafo a partire da uno o più nodi iniziali. Anche in questo caso non esistono né schemi di riferimento né linguaggi standard. Sono particolarmente adatti quando si vuole gestire una base di dati che ha una naturale rappresentazione mediante un grafo (per esempio una rete sociale). I graph databases più popolari sono Neo4j e OrientDB.

- Column stores: i dati sono raggruppati in record (row) fatti da un certo numero di campi (column) che può variare in numero e tipologia da record a record. L'accesso ai dati avviene sulla base di valori per le colonne (selezione) e degli attributi di interesse (proiezione). La rappresentazione è affine al modello relazionale (righe e colonne) ma senza alcun vincolo sulla struttura delle righe. Ogni sistema adotta un linguaggio proprietario anche se sono diffusi linguaggi SQL-like, vista la similitudine con il modello relazionale. I column store più popolari sono HBase e Cassandra.

- Key-Value databases: utilizzano un modello di rappresentazione dei dati molto semplice. Ogni dato è una semplice coppia (chiave,valore) nella quale il secondo campo può essere di grandi dimensioni. L'unica modalità di accesso offerta è sulla base di un valore per la chiave; il campo 'valore' è opaco per l'utente e non è possibile fare ricerche basate sul suo contenuto. Non esiste quindi uno schema a cui riferirsi e le primitive di accesso ai dati sono molto semplici (put e get). I key-value database più popolari sono Riak e Redis.

Le varie tipologie di sistemi NoSQL sono riassunte in Figura 5.

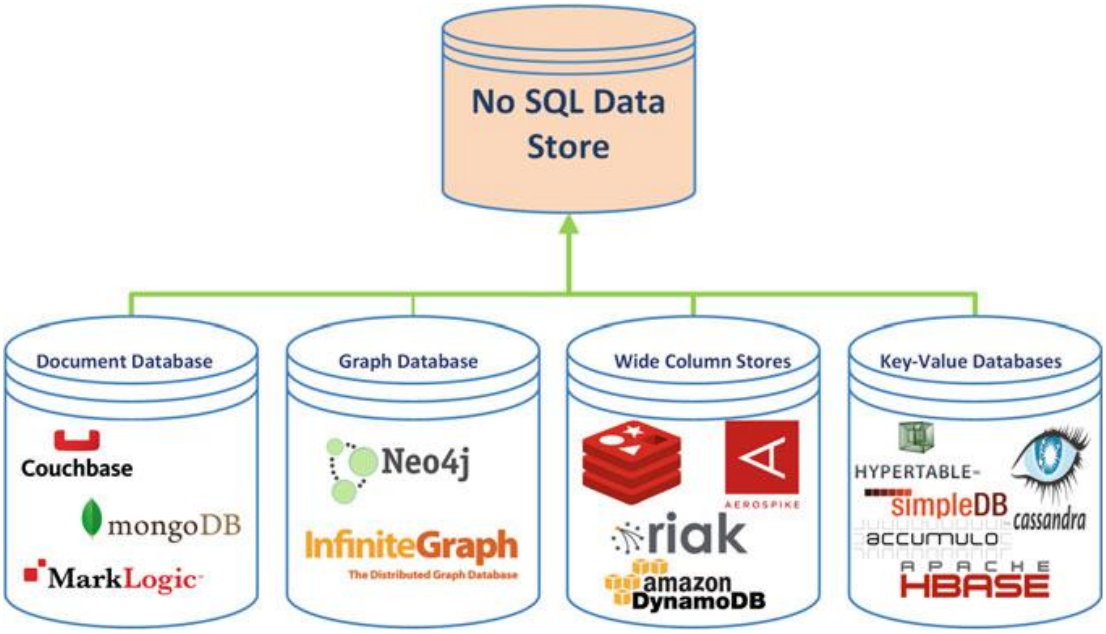


Figura 5 – Tipologie di sistemi NoSQL

Cloud computing

Il cloud computing può essere visto come una forma sofisticata di hosting, in cui un fornitore mette a disposizione tutte le risorse hardware, di rete e software necessarie e inoltre consente all'utente di gestire la propria architettura attraverso interfacce remote, tipicamente Web. Il cloud computing è quindi definibile come un modello che consente l'accesso tramite rete internet a un insieme condiviso di risorse (hardware, software di base e applicativo, connessione in rete, spazio di memoria persistente) che possono

essere allocate e rilasciate in modo semplice, senza necessità di interazione fisica con il fornitore. Un aspetto critico di questo scenario è proprio la possibilità di disporre in modo facile di un gran numero di unità di elaborazione, che possono essere sfruttate per gestire grandi quantità di dati solo grazie alla distribuzione attenta del carico di lavoro su diversi nodi.

Prima di considerare gli aspetti specifici legati al parallelismo, analizziamo le varie caratteristiche che distinguono il cloud computing dalle normali architetture dei sistemi informativi:

- *Self-service su richiesta*: gli elementi dell'architettura possono essere definiti dall'utente in base alle proprie esigenze e allocati in modalità self-service, grazie a semplici interfacce Web.

- *Accesso via rete*: la gestione dell'architettura avviene da remoto: si possono comandare la creazione e configurazione degli elementi architetturali, l'installazione delle applicazioni, la messa online e lo spegnimento.

- *Misurazione del servizio*: le risorse dell'architettura vengono affittate secondo un modello di pagamento che presuppone il monitoraggio continuo della quantità e qualità delle risorse prese in affitto, siano esse capacità di rete, macchine virtuali o licenze software.

- *Elasticità*: la configurazione dell'architettura non è immutabile, ma può essere aggiornata in qualunque momento, sia manualmente dall'utente sia da regole che monitorano il carico delle applicazioni e reagiscono alle variazioni. Per esempio, è possibile definire regole che aumentano il numero di macchine virtuali quando il numero di richieste concorrenti o i tempi di risposta superano una soglia massima.

- *Condivisione delle risorse*: le risorse della infrastruttura cloud non sono normalmente riservate all'utente ma possono essere condivise tra molte applicazioni, anche di utenti diversi. Per economia di gestione, spesso le infrastrutture cloud sono costituite da moltissime macchine fisiche tutte identiche, che vengono virtualizzate mediante software di gestione appositi in modo da emulare l'ambiente di esecuzione richiesto dal client.

Il modello del cloud computing può essere realizzato in modi diversi, a seconda di dove risiedano le risorse e del grado di condivisione (vedi Figura 6):

- *Cloud privata*: questa configurazione utilizza il modello del cloud computing per realizzare una infrastruttura proprietaria, quale l'infrastruttura di calcolo di un'impresa o di un'organizzazione. Per esempio, il CERN di Ginevra, l'istituzione dove è nato il World Wide Web, utilizza una cloud privata di grandi dimensioni e distribuita in più paesi per memorizzare ed elaborare l'enorme quantità di dati prodotti dagli esperimenti di collisione.

- *Cloud comunitaria*: l'infrastruttura di cloud appartiene a un'istituzione comunitaria, che ne garantisce l'uso agli aderenti. Un esempio è l'infrastruttura federale di cloud computing degli Stati Uniti, che mette a disposizione le proprie risorse alle agenzie federali.

- *Cloud pubblica*: l'infrastruttura cloud è esposta al pubblico generale dei clienti, che possono prendere in affitto le risorse. Tra i fornitori di servizi di cloud pubblica più usati ci sono: Amazon Web Service, Microsoft Windows Azure, Google Cloud Platform, Salesforce, Cloudbees, e Rackspace,

- *Cloud ibrida*: a questa categoria appartengono le soluzioni che fanno uso di più modelli contemporaneamente. Un caso tipico è una cloud privata che utilizzi risorse aggiuntive prese da una cloud pubblica per fare fronte a picchi di domanda.

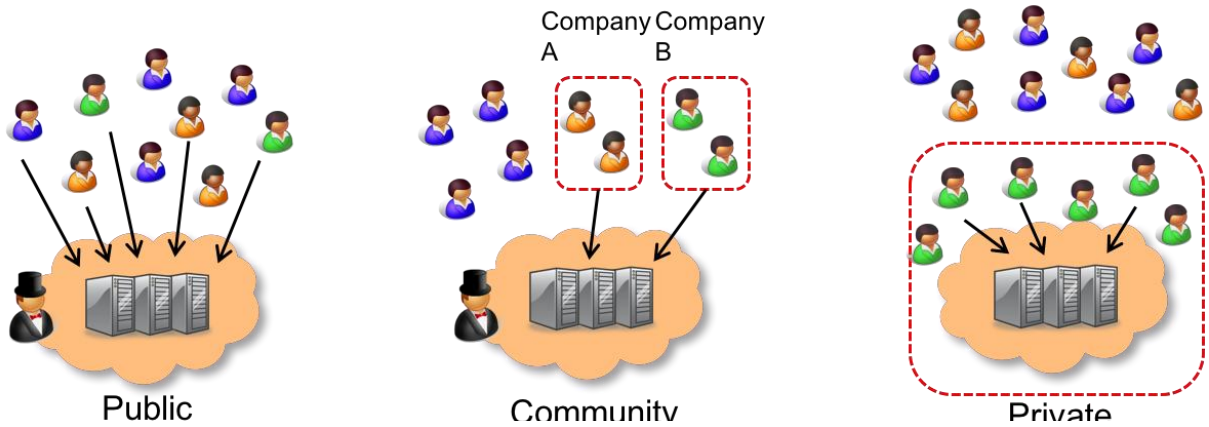


Figura 6 – Tipologie di cloud a confronto

Dal punto di vista del fornitore, un'infrastruttura cloud si presta alla realizzazione di modelli di servizio differenti.

- *Cloud Software as a Service (SaaS)*: il fornitore mette a disposizione tutto ciò che è necessario all'uso di un pacchetto applicativo. Il cliente noleggia l'uso di un'applicazione finita. Salesforce.com esemplifica questo modello di offerta e mette a disposizione dei clienti un sistema gestionale affittabile su cloud.

- *Cloud Platform as a Service (PaaS)*: il fornitore mette a disposizione le risorse hardware e il software di base. Il cliente usa la piattforma risultante per installare le proprie applicazioni. Cloudbees esemplifica questo modello di offerta e mette a disposizione dei clienti una piattaforma completa per lo sviluppo di applicazioni Web in Java.

- *Cloud Infrastructure as a Service (IaaS)*: il fornitore mette a disposizione sol- tanto le risorse hardware. Il cliente usa tali risorse per installare il software di base e le proprie applicazioni. Amazon Web Services e RackSpace esemplificano questo modello di offerta.

Altri strumenti per la gestione di big data

Oggi sono disponibili tantissimi altri strumenti per i big data oltre a quelli elencati in questa sezione. Ne indichiamo solo alcuni, che verranno citati nelle sezioni che seguono.

- Strumenti per la cosiddetta "Data Ingestion" per collezionare, aggregare e spostare da un sistema ad un altro dei big data. Per esempio, Flume, Sqoop e Kafka. In particolare, Apache Kafka (http://kafka.apache.org/) è un messagge broker, ovvero uno strumento che è in grado di gestire code di flussi di dati che arrivano in maniera asincrona da produttori esterni (per esempio una rete di sensori) e servire diversi consumatori di questi dati (per esempio un framework per la gestione di big data). Apache Sqoop (http://sqoop.apache.org/) è invece uno strumento software in grado di trasferire dati e da un DBMS relazionale ad Hadoop e popolare tabelle di Hive o HBase.

- Sistemi per lo scheduling e il coordinamento dei processi che operano su dati memorizzati in un cluster. Per esempio Zookeeper e Oozie.

- Sistemi per il "System Deployment " ovvero per la gestione dell'infrastruttura hardware (cluster management). Per esempio Ambari e Mesos.

2.1.2 Una architettura big data per lo smart street

Vediamo ora come le tecnologie per i big data appena discusse possono essere utilizzate efficacemente nello scenario di interesse dello smart ligthing.

Innanzitutto va osservato che, che per quanto detto sopra, l'infrastruttura di base è un cluster di computer accessibile inizialmente su cloud privata. Questo consentirebbe la sperimentazione e il test in ambiente chiuso e garantirebbe un buon livello di sicurezza. Successivamente si può studiare una soluzione di cloud ibrida che utilizzi una cloud pubblica (per esempio quella offerta da Amazon o Microsoft) se l'infrastruttura a disposizione non fosse in grado di supportare la crescita del volume dei dati.

Tenendo conto poi delle caratteristiche dell'applicazione di riferimento nella quale sono necessarie sia attività di monitoraggio dei dati in real-time, sia l'analisi di dati raccolti in ampi intervalli temporali per comprendere andamenti e criticità una possibile soluzione tecnologica a supporto della raccolta e l'analisi dei dati generati da smart street è la cosiddetta architettura "lambda".

Come illustrato in Figura 7, questa architettura è dotata delle seguenti componenti principali:

- Un **data-broker layer** in grado di acquisire e gestire code di stream di dati generati in tempo reale da sistemi di sensoristica remota e rifornire con questi dati diversi sistemi di elaborazione. Per questa componente si fa tipicamente uso di un message broker per big data come Kafka.

- Un **Batch layer** per la memorizzazione incrementale dei dati grezzi che non verranno mai modificati (immutable data) in modalità "append-only" e per le analisi su intervalli temporali ampi. Il repository che man mano si crea viene in genere chiamato "master dataset". Per realizzare il batch layer si fa uso tipicamente di un file system distribuito come HDFS e di strumenti di elaborazione parallela per Hadoop come MapReduce, Spark, Hive e Pig. I dati e le analisi estratte dal master dataset vengono tipicamente memorizzati su viste materializzate memorizzate in maniera persistente, ancora su Hadoop o su sistemi NoSQL. Le analisi di output vengono generate da queste viste.

- Un **Realtime layer** per la memorizzazione e l'analisi dei flussi di dati in tempo reale su intervalli temporali brevi e recenti. Spesso i dati raccolti devono essere integrati con i nuovi dati per aggregare dati puntuali e creare intervalli temporali di varia ampiezza. Le tecnologie utilizzate per questa componente includono sistemi di stream processing quali Flink, Storm e Spark Streaming.

- Un **Serving layer** che memorizza in maniera persistente i risultati generati dal batch layer e dallo speed layer sotto forma di viste materializzate. Per questa componente si fa in genere uso di strumenti di database management NoSQL come HBase, MongoDB e Cassandra.

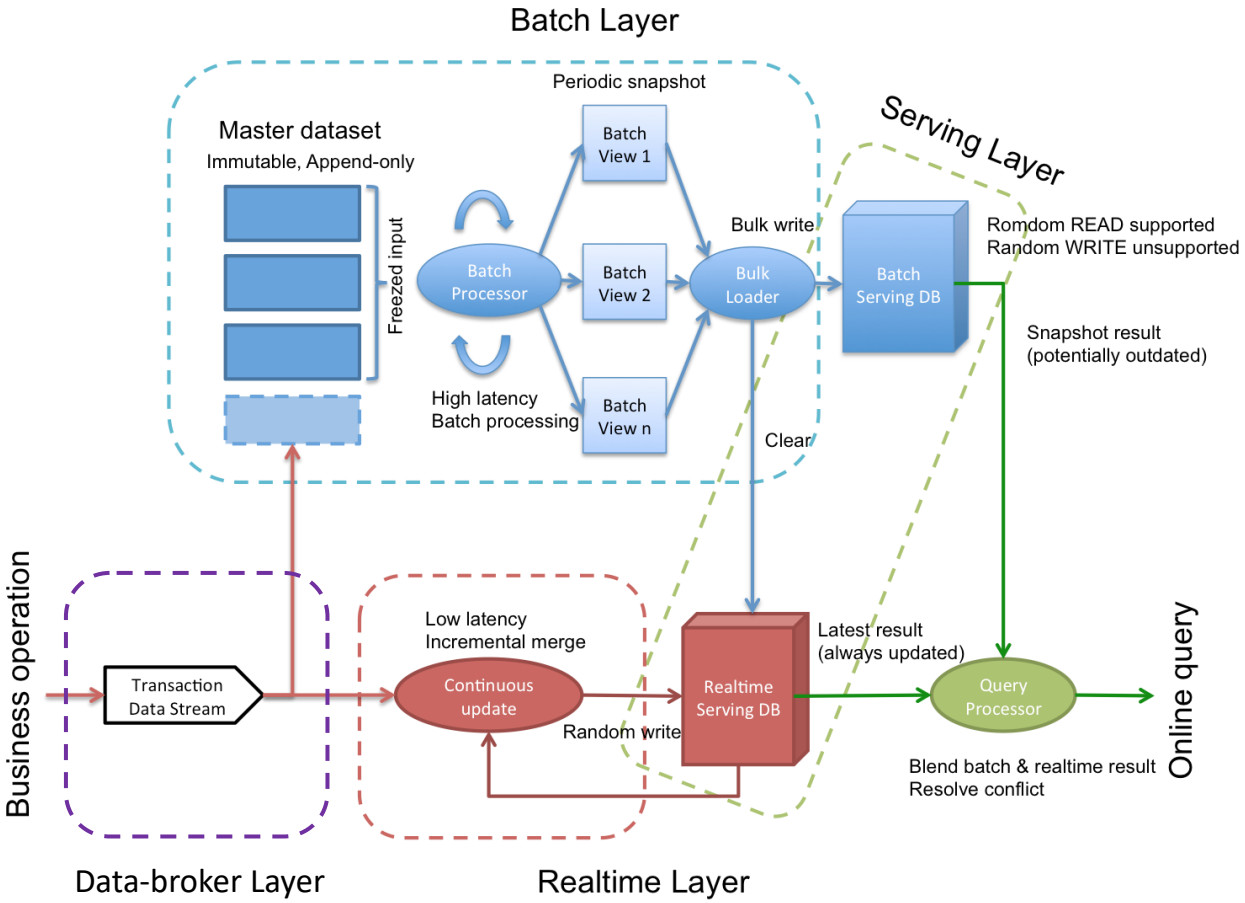


Figura 7 – l'architettura lambda per l'analisi dei big data.

Una possibile alternativa all'architettura lambda e la cosiddetta architettura kappa, che consiste in una semplificazione dell'architettura lambda ottenuta realizzando il batch layer per mezzo del Realtime layer stesso. Questa idea si basa sulla osservazione il batch layer accumula dati che hanno un istante di inizio e uno di fine mentre un stream di dati è concettualmente infinito, non avendo un inizio e una fine. Se ne conclude che il dataset del batch layer è sempre un sottoinsieme del dataset del realtime layer . Quindi, per realizzare il batch layer è sufficiente sfruttare i dati raccolti dal realtime layer una volta che questi sono stati opportunamente accumulati su un (master) data set unificato.

In Figura 8 è mostrato un confronto tra le due architetture.

La scelta tra architettura lambda e architettura kappa dipende da diverse questioni che includono i vincoli implementativi, i casi di uso del sistema e i requisiti di efficacia ed efficienza. Generalmente, si parte da un'architettura lambda che separa i due ambienti batch e realtime e successivamente si verifica se è possibile procedere ad una semplificazione architetturale che non peggiori le performance.

Una volta realizzata e testata l'architettura del sistema, è anche necessario individuare un processo di migrazione di dati, viste e interrogazioni definite sul sistema preesistente nella nuova piattaforma. Per questa sistema è possibile utilizzare applicativi ad-hoc come Sqoop.

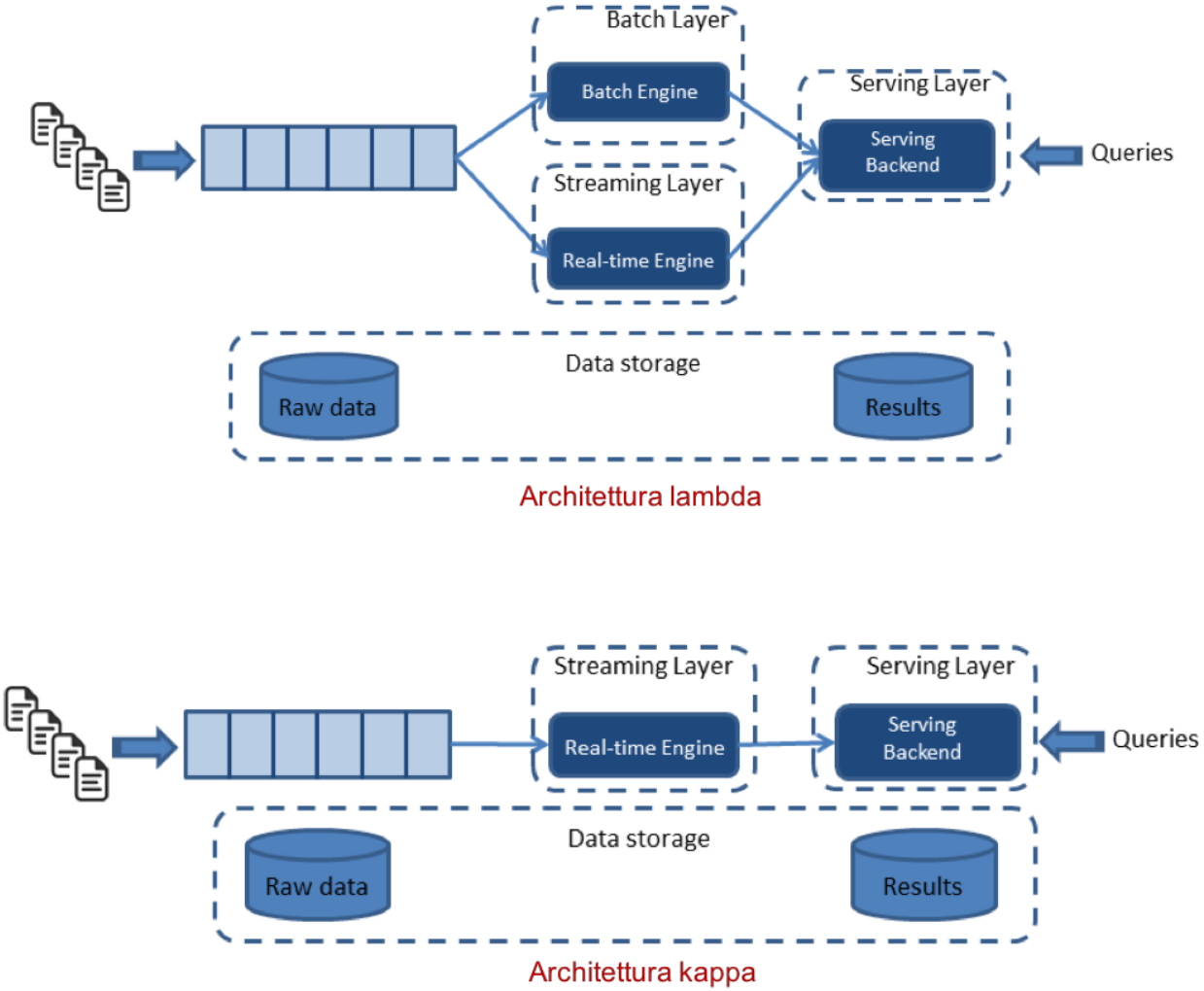


Figura 8 – Architettura lambda e kappa a confronto

2.1.3 Tipologie di analisi

Ci occupiamo ora di mostrare come l'architettura proposta nella precedente sezione possa essere messa a servizio delle esigenze specifiche di una applicazione di smart street/ligthing e, più in generale, di una applicazione per l'IoT.

In questi contesti, l'obiettivo principale è quello di offrire degli strumenti efficaci ed efficienti per l'analisi delle misurazioni fatte da una rete di sensori. Questo problema pone diverse sfide, a causa della tipologia dei dati raccolti e della quantità e rapidità con cui questi dati vengono raccolti nel corso del tempo.

Un tipico problema nell'analisi sensoristica è quello di rilevare eventi in tempo reale. In questo contesto la sfida è che gli eventi semantici di alto livello sono spesso una funzione complessa dei dati del sensore grezzi sottostanti. In alcuni casi, l'evento originale non può essere rilevato con esattezza, poiché il processo di rilevamento dell'evento è correlato ai dati in modo ambiguo. Inoltre, i dati di misurazione dei sensori sono intrinsecamente rumorosi ed incerti, e potrebbero essere affetti da misurazioni mancanti o ridondanti, a seconda del dominio. Oltre alla rilevazione di eventi real-time, l'utente può essere interessato a considerazioni globali, per esempio sul consumo energetico globale o medio in certi intervalli temporali, sull'efficienza delle apparecchiature, o sui macro-trend del fenomeno fisico osservato. Queste analisi richiedono l'elaborazione di collezioni di dati più grandi rispetto alla rilevazione di eventi e producono in genere della reportistica che non richiede vincoli temporali stringenti per poter essere prodotta.

In generale, è possibile quindi individuare due tipologie di analisi, cosiddette *online* e *offline*, da applicare rispettivamente per rilevare eventi e per effettuare delle considerazioni globali.

- L'analisi offline (o "batch") offre una vista sull'intero dataset a disposizione o su sue grossi porzioni. Di conseguenza la frequenza di aggiornamento non può essere molto alta: all'arrivo di nuovi dati essi sono inseriti nello storage e aggregati in viste specifiche per le analisi di interesse, che vengono ricalcolate da zero. Le operazioni di aggregazione posso sono essere efficacemente implementate con paradigmi di elaborazione parallela, per esempio, con job MapReduce o Spark.

- L'analisi online (o "realtime") compensa l'alta latenza del livello offline, basandosi solo sui nuovi dati ricevuti: all'arrivo di nuovi dati, essi vengono usati per incrementare le viste esistenti con il nuovo contenuto. Il risultato del calcolo ha di conseguenza natura transitoria. Le operazioni di incremento sono implementate con paradigmi di data streaming, per esempio, con Spark Streaming o strumenti analoghi.

È ben evidente che l'architettura presentata nella sezione precedente si adatta bene a questi requisiti di analisi: il batch layer sarà a servizio dell'analisi offline mentre il realtime layer a quello dell'analisi online.

2.1.4 Analisi batch

Le tipiche analisi batch mirano a calcolare statistiche globali sui dati di misurazione di sensori e richiedono di aggregare i dati in base a due problemi distinti:

- Integrazione temporale: identificare quali fenomeni non cambiano significativamente nel tempo [13]
- Integrazione spaziale: identificare quali sensori hanno rilevato fenomeni simili [14]

La sfida maggiore in entrambi i casi è quella di implementare efficientemente il calcolo, dato il grande volume di dati e di sensori coinvolti. I paradigmi di calcolo più diffusi in questo ambito sono MapReduce e Spark.

Di seguito riportiamo alcune tipologie di analisi batch, con riferimento alle grandezze rilevabili nel contesto dello smart lighting e, più in generale, nel contesto dell'IoT per il monitoraggio di una smart city.

- Livello di illuminazione:
  - livello di illuminamento medio per data fascia oraria,
  - giorni in cui in certe fasce orarie il livello di illuminamento è diverso dalla media.
- Dati meteo ed energia termica prodotta
- Consumo elettrico illuminazione
- Consumo elettrico emergenza
- Consumi Accensione contemporanea di un numero anomalo di utenze elettriche rispetto al livello di occupazione (F.E.M.)
- Consumo elettrico fan coil
- Velocità ventole fan coil
- Consumo elettrico globale
- Portata in volume fluido raffreddamento
- Energia termica caldaia
- Portata in volume fluido riscaldamento
- Energia termica effettiva su energia termica erogata

- Temperatura interna edificio
- Umidità interna edificio
- Misura CO2
- Afflusso dipendenti

2.1.5 Analisi real-time

Le tecniche tipiche per rilevare eventi in dati di misurazione di sensori, sono più in generale tecniche di rilevazione degli outlier. Nel problema generale di rilevazione degli outlier, l'obiettivo è quello di identificare i valori di misurazione che sembrano molto diversi dai loro vicini spazio-temporali, ovvero, i valori della storia recente del flusso del sensore, oppure i valori dei flussi di sensori spazialmente vicini. Specifichiamo che l'evento rilevato può o essere rappresentativo di un evento del mondo reale, e dunque di interesse per se, o prodotto dal rumore, e quindi da pulire. Diversi metodi sono orientati o all'una o all'altra applicazione.

I metodi proposti in letteratura sono di tre tipi:

- Classification-based: Nei metodi basati su classificatori (es., [3,4]) si assume che la lettura attuale di ciascun sensore venga influenzata solo dalla lettura precedente dello stesso sensore, e le letture dei suoi vicini. Questo permette di calcolare la distribuzione di probabilità associata al sensore e decidere se la lettura attuale è attesa o meno, classificandola come outlier di conseguenza. I metodi in letteratura si differenziano per il classificatore usato (es., bayesiano in [3] e SVM in [4]) e per le specifiche assunzioni nel modello.

- Data Distribution-based: Nei metodi basati su distribuzione dei dati (es., [5,6,7,8]) si apprendono, per ciascun sensore, le distribuzioni delle differenze delle sue letture nel tempo e delle letture dei suoi vicini. Il principio è simile a quello dei metodi basati su classificatori, ma non è richiesta conoscenza preliminare delle distribuzioni di dati. Ad esempio, [7] identifica quei valori per cui sono stati letti pochi valori simili (cosiddette anomalie basate sulla distanza), mentre [8] seleziona quei valori nelle regioni più sparse dello spazio dei dati (cosiddette anomalie basate sulla densità). Questo richiede solo di contare il numero di valori in diverse regioni dello spazio.

- Node Similarity-based: Nei metodi basati su similarità dei nodi (es., [9,10]) l'obiettivo è individuare e pulire i valori anomali. [9] si concentra su due tipi di letture anomale: brevi o lunghe. Ad alto livello, l'approccio descritto calcola una forma di trasformata wavelet discreta (DWT) sulla serie di letture di ogni sensore. Le letture che sono sufficientemente lontane dalla trasformata, vengono classificati come anomalie brevi. Le intere serie che sono sufficientemente lontane da quelle degli altri sensori vicini (secondo la cosiddetta dynamic time warping distance) vengono classificate come anomalie lunghe. [10] calcola, per ogni sensore, la differenza tra sua lettura e la lettura mediana dei sensori vicini. Se le differenze di un dato sensore superano una soglia, allora il sensore viene classificato come anomalo.

Al fine di ottimizzare l'implementazione di tali metodi, consideriamo tecniche di data streaming [2]. Tali tecniche forniscono buone approssimazioni di molte funzioni di aggregazione standard (es. frequency moments). Poiché Interpretano i dati di misurazioni di sensori come generiche serie multidimensionali, non richiedono alcuna conoscenza a priori del processo fisco osservato. Inoltre, queste tecniche possono essere impiegate efficientemente su architetture distribuite.

2.1.6 Data cleaning

La qualità dei dati è una componente importante del processo di analisi. I dati di misurazione dei sensori sono intrinsecamente rumorosi ed incerti, a causa, per esempio, di eventuali guasti di rete, batterie scariche, congelamento o il riscaldamento del dispositivo di custodia o di misura, accumulo di sporcizia, o guasti meccanici. Ciò può causare un problema significativo per quanto riguarda l'utilizzo dei dati, dal momento che le applicazioni che utilizzano i dati errati possono portare a risultati non accurati. Per risolvere questo problema, è essenziale rilevare e correggere i valori errati nei dati di misurazione dei sensori.

In caso di dati rumorosi, un approccio tipico è quello di pulire i dati basandosi su modelli consolidati (es., [11]). In questo approccio, i valori del sensore più probabili sono desunti dai modelli, e le anomalie vengono rilevate confrontando i valori dei sensori osservati con i corrispondenti valori desunti. Anche se la qualità delle misurazioni riportate dai sensori è garantita, l'identificazione dei valori anomali fornisce un modo efficiente per guidare la rilevazione di eventi interessanti. In caso di dati mancanti, è possibile applicare tecniche di interpolazione delle misurazioni dei sensori (questo richiede la conoscenza fisica dei processi coinvolti).

2.1.7 Migrazione nel nuovo sistema

L'attività di migrazione dei dati nella nuova piattaforma deve prevedere il mantenimento dei dati raccolti e delle principali funzionalità di analisi già in essere. Visto che attualmente i dati sone memorizzati su un sistema relazionale e le procedure di accesso ai dati sono state implementate in SQL, in prima battuta si può pensare di effettuare un trasferimento di dati e operazioni che:

- preservi per quanto possibile la natura relazionale dei dati anche nel momento in cui verranno riversati su Hadoop,

- utilizzi per l'implementazione delle procedure esistenti che si occupano di estrarre i dati di interesse una soluzione SQL su Hadoop.

Un processo di migrazione può quindi essere riassunto come segue:

1. Trasferimento dei dati su Hadoop. A questo scopo si può usare Sqoop che consente di trasferire dati da una base di dati relazionale su Hadoop. Questo strumento supporta caricamenti incrementali di una singola tabella o query SQL, o job salvati che possono essere eseguiti più volte per importare aggiornamenti fatti alla base di dati rispetto dall'ultimo trasferimento. Può importare i dati direttamente in Hbase o Hive. Inoltre è possibile eseguire il processo inverso: da Hadoop a tabelle relazionali.

2. Definizioni di viste che ricostruiscono nel batch layer o nello speed layer le query originari di tipo offline e online rispettivamente utilizzando un linguaggio SQL-based, per esempio quello di Hive o HBase.

3. Test di funzionamento delle procedure ricostruite e definizione di procedure di tuning per la loro ottimizzazione su Hadoop.

4. Realizzazione di nuove funzionalità di analisi utilizzando strumenti quali MapReduce e Spark sul batch layer e Spark streaming sullo speed layer.

2.2 Cyber security in uno scenario Smart Street

2.2.1 Inquadramento

Per la criticità delle comunicazioni nello sviluppo di sistemi di SL (e più in generale di SC), sono state individuate alcune specifiche, utili alla definizione degli obiettivi del presente lavoro, che possono essere riassunte come segue [1]:

- *Alto livello di sicurezza*: le comunicazioni devono essere sicure dal punto di vista cyber [2]. I rischi informatici da considerare hanno una duplice natura. Bisogna considerare i rischi dovuti ad attacchi informatici deliberati verso il sistema [3] e quelli dovuti a malfunzionamenti, disastri naturali o alterazione inavvertita del flusso informativo;

- *Qualità del Servizio (QoS)*: le comunicazioni tra gestore dell'infrastruttura e utente sono fondamentali per garantire un servizio di qualità [4]. I dati relativi al funzionamento del sistema (dati di monitoraggio, risposta agli allarmi, comandi di controllo, etc.) devono essere trasferiti entro una determinata finestra temporale, senza che la rete di comunicazione venga sovraccaricata;

- *Privacy/Confidenzialità*: i dati sensibili che viaggiano nei sistemi devono essere protetti [5] e gli storici di funzionamento del sistema non dovrebbero costituire sistemi per scoprire le abitudini dei singoli utenti;

- *Integrità*: si riferisce all'affidabilità dei dati e delle risorse. Da un punto di vista operativo si traduce nella capacità di prevenire cambiamenti nei dati impropri o non autorizzati. L'integrità è un aspetto cruciale nella progettazione di sistemi informativi e di sistemi di comunicazioni [6]. Nei sistemi di SL si applica alle informazioni quali gli algoritmi di controllo, i valori dei sensori e i valori dei comandi di controllo. L'integrità viene raggiunta difendendo il sistema da attacchi volti a modificare messaggi (*message injection, message replay, message delay* etc.). La violazione dell'integrità può portare a problemi di sicurezza [7].

- *Affidabilità e continuità dei servizi*: l'affidabilità è sicuramente un requisito fondamentale nei sistemi di illuminazione pubblica. La diffusione di protocolli di comunicazione sicuri e lo sviluppo dei sistemi informatici ed embedded ha portato ad una maggiore affidabilità dei dispositivi intelligenti. Per ottenere la continuità nella disponibilità del servizio, il sistema di comunicazione gioca un ruolo fondamentale: per garantire il flusso di informazioni vengono pertanto scelti sistemi ibridi basati su diversi canali di comunicazione.

I sistemi di comunicazione rappresentano quindi un anello fondamentale nella realizzazione di un sistema SL. Questi sistemi, per la loro natura, sono anche i sistemi più soggetti ad attacchi. Il presente task si pone l'obiettivo di analizzare la sicurezza dei sistemi di comunicazione delle SL con particolare riferimento alla Smart Street installata presso la sede di ENEA – Casaccia. A tal fine, si analizzeranno in maniera generale i sistemi di comunicazione delle SL, considerando sia reti *wired* basate sulle onde convogliate, ossia Power Line Communications (PLC) che reti wireless basate su reti WiFi o reti ad hoc. Successivamente, verranno introdotte le principali vulnerabilità dei sistemi in oggetto. Infine, verrà descritto il sistema Smart Street ENEA e ne verrà analizzata la sicurezza.

L'obiettivo è quello di fornire un quadro sulle possibili vulnerabilità del sistema sulla base delle scelte progettuali. I limiti delle analisi presentate sono dati dalla metodologia utilizzata. Non si è potuto sottoporre il sistema in oggetto a stress-test sistematici che potessero mettere in evidenza specifiche criticità. L'analisi proposta, pertanto, si basa su una revisione delle scelte progettuali effettuate ed una valutazione delle soluzioni o le mitigazioni adottate rispetto alle vulnerabilità precedentemente analizzate.

2.2.2 Power Line Communication

La tecnologia Power Line Communications (PLC) [8] utilizza le linee di distribuzione dell'elettricità per creare un'infrastruttura di comunicazione su cui basare alcuni servizi.

Questo tipo di tecnologia di comunicazione porta diversi vantaggi sia per i gestori delle linee di distribuzione elettrica che per gli utenti finali. Dal punto di vista dei gestori, questi hanno a disposizione un'unica infrastruttura per fornire servizi legati alla distribuzione elettrica e alla comunicazione. Questo porta ad un'evidente riduzione dei costi finali di esercizio e degli sprechi per la realizzazione di infrastrutture diverse. Dal punto di vista degli utenti finali, questi possono accedere ad una maggiore offerta di servizi come il controllo remoto dei consumi di un'abitazione di un palazzo, la gestione intelligente dei dispositivi presenti. Reti PLC locali, infatti, sono ampiamente utilizzate nella *building automation* per applicazioni si sicurezza, supervisione del condizionamento, controllo dell'illuminazione.

Al momento, ci sono diverse attività connesse con lo sviluppo di applicazioni per le PLC. Le reti di distribuzione ad alta e media tensione, infatti, possono essere utilizzate come ponti per coprire grandi distanze ed evitare la costruzione di nuove reti di telecomunicazione. Le reti a bassa tensione sono disponibili in grande numero e possono essere utilizzate per realizzare l'accesso alla rete PLC ed evitare il noto problema dell'*ultimo miglio*.

Rete elettriche per le PLC

I sistemi di trasmissione e distribuzione elettrica consistono di tre livelli di rete che possono essere usate come mezzo di trasmissione per la realizzazione di reti PLC:

- *Rete ad alta tensione* (110-380 kV) rappresenta la rete di trasmissione del sistema elettrico: è il passaggio intermedio tra produzione e la distribuzione dell'energia elettrica. Essa connette le stazioni ricevitrici (*power station*) di grandi regioni e copre solitamente grandi distanze, consentendo lo scambio di energia elettrica tra zone molto distanti. Le reti ad alta tensione sono realizzate con cavi aerei.

- *Rete a media tensione* (10-30kV) distribuisce energia elettrica a grandi aree, città e utenze industriali o commerciali. Essa copre distanze inferiori rispetto alle linee ad alta tensione ed è realizzata sia attraverso linee aeree che interrate.

- *Rete a bassa tensione* (230/400 V o 110 V) fornisce corrente agli utenti finali, sia come consumatori finali. La rete a bassa tensione copre poche centinaia di metri. Nelle aree urbane, la rete a bassa tensione è realizzata con cavi interrati, mentre nelle aree rurali ancora sussistono linee aeree.

Le installazioni elettriche all'interno delle abitazioni appartengono alla rete a basso livello, ma sono proprietà dell'utente. Le utenze private sono connesse alla rete di distribuzione mediante un contatore. Il resto della rete elettrica appartiene alle società di distribuzione elettrica. Le reti a bassa tensione sono ampiamente diffuse e interagiscono con l'utente finale: questa caratteristica rende più semplice lo sviluppo di applicazioni basate su PLC e reti a bassa tensione. La rete a bassa tensione, tuttavia, copre solo poche centinaia di metri tra i consumatori e le unità di trasformazione offrono un'alternativa all'utilizzo di PLC per la realizzazione dell'ultimo miglio nelle reti di telecomunicazione.

Trasmissione del segnale

La trasmissione di informazioni su reti di trasmissione o distribuzione può essere divisa in due tipologie in base alla tecnologia con cui la PLC viene implementata. Queste due tipologia operano a bande diverse ed hanno diversa capacità.

I sistemi PLC possono essere:

- *Narrow Band* (NB-PLC): consentono servizi di comunicazione a bassa velocità (fino a 100 kbps) e assicurano la realizzazione di applicazioni di automazione e controllo insieme a pochi canali voce;

- *Broad Band* (BB-PLC): consentono servizi di comunicazione ad alta velocità (oltre 200 Mbps) e possono realizzare in maniera parallela diversi servizi, quali l'accesso a Internet e la telefonia.

Le NB-PLC operano ad una frequenza inferiore a 500 kHz: le frequenze fino a 148.5 kHz sono state assegnate dal CENELEC (l'ente europeo per le standardizzazioni) per i sistemi NB-PLC per le comunicazioni su reti di distribuzione elettrica. In questo insieme di frequenze la velocità di trasmissione è modesta: essa è compresa tra 1 Kbps e 100 Kbps. Queste velocità di trasmissione sono, tuttavia, adeguate per applicazioni di controllo e monitoraggio remoto. Nell'America settentrionale, in Giappone e in Cina, la frequenza assegnata dagli enti di standardizzazione arriva fino a 500 kHz, rendendo possibile una capacità di comunicazione più grande (300 kbps), che consente l'implementazione di semplici servizi.
Servizi complessi, che richiedono una maggiore banda, non possono essere implementati e richiedono la tecnologia BB-PLC.

In aggiunta alle NB-PLC, altre soluzioni implementative, relative alla tecnologia BB-PLC, sono utilizzate comunemente in ambienti domestici. Questi sistemi sono in grado di fornire dati ad alta velocità, fino a diverse centinaia di Mbps, ma sono più sensibili ai rumori sulla rete elettrica, limitando la loro applicazione in ambiente industriale.
Le BB-PLC operano nelle bande di frequenza HF/VHF (1.8-250 MHz), permettendo di raggiungere velocità di trasmissione di 200 Mbps, utilizzando modulazioni BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM e 1024-QAM nello schema OFDM.

Le reti di distribuzione elettrica a bassa tensione di tipo BB-PLC sembrano essere la soluzione più efficiente da un punto di vista costo-beneficio per risolvere il problema dell'ultimo miglio nelle reti di comunicazione.
L'applicazione delle reti elettriche di distribuzione elettrica nelle telecomunicazioni sono note fin dall'inizio del ventesimo secolo. Il primo *carrier frequency system* ha operato in una rete di distribuzione elettrica ad alta tensione in grado di coprire una distanza di oltre 500 km con un segnale a 10 W [9]. Questi sistemi sono stati utilizzati per realizzare sistemi di misura e controllo remoti. Anche le comunicazioni su reti a media o bassa tensione sono state realizzate. Sistemi basati su *ripple carrier signaling* sono stati applicati a reti di bassa e media tensione per la gestione dei carichi elettrici nelle reti di distribuzione.

Il canale di comunicazione realizzato tramite rete elettrica presenta delle caratteristiche particolare per l'alta rumorosità, la presenza di multipath e la selettività del canale stesso. Queste caratteristiche non consentono una corretta trasmissione delle informazioni se queste non vengono opportunamente codificate mediante modulazione. La scelta della tecnica di modulazione, pertanto, dipende fortemente dal canale e, nel caso delle reti elettriche, deve essere in grado di contrastare le caratteristiche non lineari del canale. Quest'ultime, infatti, potrebbero rendere la demodulazione molto complessa e molto costosa, se non impossibile, per velocità superiori a 10 Mbps utilizzando una modulazione a singolo canale (*Single carrier modulation*). Quindi, la modulazione per PLC deve mitigare questo problema introducendo equalizzazioni complesse.

I cambiamenti di impedenza nella rete elettrica portano ad eco che causano dispersioni e ritardi: la modulazione deve pertanto riuscire a gestire i multipath che si vengono a creare. La scelta della modulazione deve offrire alta flessibilità e/o la capacità di evitare alcune specifiche frequenze, se queste sono disturbate oppure allocate ad altri servizi.
In questi ambienti una modulazione ha mostrato buone performance in ambienti difficili e, quindi, sono state largamente diffuse e sviluppate. Questa è la OFDM (*Orthogonal Frequency Division Multiplexing*), già adottata dalla European Digital Audio Broadcasting (DAB).

Il principio di funzionamento della *MultiCarrier Modulation* (MCM) si basa sulla divisione del flusso di dati da trasferire in flussi di bit paralleli, ognuno avente un bit rate minore. Ogni flusso utilizza diverse portanti, dette sottoportanti per la modulazione.
Il sistema *Orthogonal Frequency Division Multiplexing* è una particolare forma di MCM avente sottoportanti densamente spaziate e spettri con sovrapposizioni. Per garantire una ricezione priva di errori dei segnali OFDM, le forme d'onda delle sottoportanti sono scelte in maniera da essere ortogonali tra loro.

Dispositivi per la realizzazione di una PLC

La rete di comunicazione basata su linee elettriche si interfaccia a diversi livelli con l'infrastruttura di comunicazione. In particolare, i dispostivi utilizzati per realizzare questo collegamento sono

- Base station;
- Modem PLC.

La *base station* ha il compito di interfacciare la rete elettrica usata per le comunicazioni con il *backbone* dell'infrastruttura di comunicazione. Una *base station*, pertanto, è dotata di una o più interfacce fisiche in grado di collegarsi con la rete PLC ed una o più interfacce specifiche verso l'infrastruttura di telecomunicazione.

Il *modem*, generalmente, è fornito di diverse interfacce per collegarsi con differenti tecnologie di comunicazione. In questo modo è possibile, ad esempio, collegarsi ad Internet con un PC tramite interfaccia Ethernet. Il *modem* è connesso anche alla rete di distribuzione elettrica con una specifica interfaccia PLC.

La base station e il modem non sono semplici convertitori di segnale, ma hanno una struttura più complessa, secondo l'architettura di rete della pila ISO/OSI [10]. Questa architettura divide la comunicazione in 7 livelli di astrazione, partendo dal supporto fisico e arrivando all'applicazione utente, come si può vedere in Figura 9.

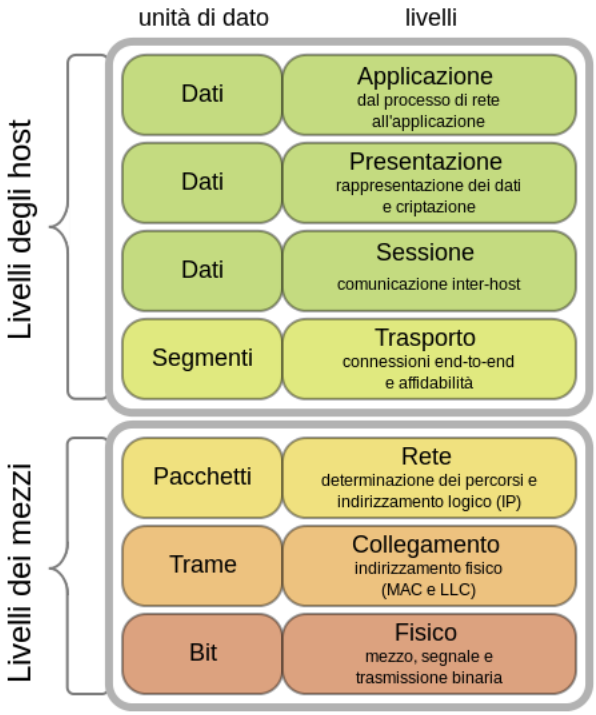


Figura 9 Modello ISO/OSI

I dispositivi in oggetto implementano i primi 3 livelli di questa architettura: il livello fisico e quello di data link è dato dall'interfaccia fisica, mentre il livello di rete è implementato via software. La conversione del segnale non viene fatta a livello fisico come nei convertitori di segnale, ma a livello di rete. Le informazioni ricevute dal livello fisico del PLC sono riportate tramite il livello LLC e MAC a livello di rete, che è organizzato secondo protocolli specifici (ad esempio, IP) che assicurano la corretta comunicazione tra PLC - Ethernet (o qualsiasi altra interfaccia).

Standard e protocolli

I sistemi PLC utilizzano diversi standard e protocolli [11] che differiscono per modulazione, frequenza e procedura di accesso al canale. I principali protocolli e gli standard sono riassunti nella Tabella 1 e nella Tabella 2 rispettivamente, in seguito viene introdotto TALQ, un protocollo specifico per le Smart Lightning.

X10	X10 è il più vecchio protocollo per PLC ed utilizza la modulazione *amplitude shift keying*. I prodotti basati su X10 ammettono comunicazione unidirezionale, pochi bidirezionale. La capacità di trasmissione e l'affidabilità non è alta su canali rumorosi. La massima banda è 60 bps su rete elettrica a 60 Hz.
CEBus	Lo standard CEBus usa la modulazione a spettro espanso. Tale modulazione parte ad una determinata frequenza e la modifica durante il suo ciclo. La velocità di trasmissione del protocollo CEBus data rate è circa 10 kbps e può essere usata per applicazioni in abitazioni o uffici. Per l'accesso al mezzo utilizza *carrier sense multiple access* (CSMA) e *collision resolution and collision detection*.
LonWorks	LonWorks è un protocollo di comunicazione peer-to-peer sviluppato dalla Echolon Corporation. L'accesso al mezzo è realizzato mediante CSMA. Lo schema di modulazione è narrowband spread-spectrum, capace di limitare gli effetti dovuti ad alta rumorosità. Esso è usato in America Settentrionale e in Europa per applicazione di controllo e automazione

	(heating, ventilating and air conditioning - HVAC).
HomePlug	Il protocollo HomePlug divide la banda in sotto-canali. La sua velocita di trasmissione varia da 1 a 14 Mbps ed i nodi scelgono in maniera automatica la velocità ottimale. L'accesso al mezzo utilizza lo schema CSMA/CD (Collision Avoidance). Per gestire la qualità del servizio, il protocollo definisce 4 livelli di priorità [12].
SCP	Il Simple Control Protocol è stato creato dalla Microsoft Corporation per fornire comunicazioni peer-to-peer sicure e robuste tra dispositivi piccoli ed economici. È adatto per reti PLC a bassa velocità e per diversi tipi di applicazioni come le reti di sensori e i sistemi di monitoraggio energetico.
LNCP	Living Network Control Protocol usa comunicazione peer-to-peer e sistemi multi-master. Il protocollo utilizza un solo canale. LNCP consente il controllo remoto dei dispositivi presenti in casa. È basato su microprocessori a 8-bit and con poca memoria.
HNCP	Home Network Control Protocol è un protocollo a 4 livelli per il monitoraggio dei dispositivi domestici: l'interoperabilità è garantita da un insieme di messaggi standard. HNCP definisce le interfacce dei modem e incapsula i suoi livelli per aumentare la robustezza. Utilizza anche un indirizzamento proprietario. I messaggi standard sono costituiti da un codice messaggio, numero di pacchetti in input e output. I nodi master possono controllare gli altri dispositivi utilizzando le interfacce specificate dallo HNCP [13].

Tabella 1 Protocolli PL

HomePlug AV	HomePlug AV è lo standard proposto da HomePlug Powerline Alliance per migliorare la versione HomePlug 1.0. HomePlug AV considera la trasmissione di contenuti audio-video all'interno di abitazioni. HomePlug AV gestisce gli errori usando codici a convoluzione. HomePlug AV utilizza la criptazione 128 bit AES.
HomePlug AV 2	HomePlug AV 2 (HomePlug Powerline Alliance) recepisce lo standard IEEE P.1901 migliorando la velocità la trasmissione. Il livello di applicazione migliora le performance rispetto alla precedente versione, consentendo più stream video simultanei. I nodi possono essere utilizzati come ripetitori per estendere la connessione in tutta l'abitazione.
HD-PLC	High Definition-Power Line Communication è uno standard definite da Panasonic ed utilizza OFDM. Per alte frequenze, HD-PLC sfrutta la modulazione Wavelet-OFDM. La massima velocità è di 210 Mbps. HD-PLC è in grado di valutare il rapporto segnale rumore del canale e adattare la velocità di trasmissione. HD-PLC è compatibile con le applicazioni di home automation e garantisce sicurezza mediante criptazione AES a 128 bit.
IEEE P.1901	Lo standard IEEE P.1901 definisce il *medium access control* (MAC) e il livello fisico. Esso utilizza frequenze inferiori ai 100 Mhz e supporta la connessione dal primo all'ultimo miglio dei servizi broadband.
ITU-T G.9960	ITU-T G.9960 è lo standard sviluppato dalla International Telecommunication Union per (ITU) per gestire le reti ad alta velocità necessarie per la HDTV (High Definition TV). Lo standard include le specifiche del livello fisico per trasmettere contenuti multimediali su cavo telefonico, coassiale e reti elettriche.
CENELEC EN 50065	CELENEC EN 50065 è lo standard di riferimento dell'Unione Europea per i sistemi PLC. Utilizza frequenze nella banda 9-140 kHz, per Stati Uniti e Giappone considera anche 500 Hz. Le applicazioni cui si riferisce sono il bilanciamento dei carichi e la misurazione remota dei consumi.
ETSI	ETSI sviluppa uno standard per l'interoperabilità tra dispositivi di diversi produttori. Il gruppo di lavoro PLT (Power Line Telecommunications) è deputato alla standardizzazione dei sistemi PLC. Esso ha prodotto diversi standard in accordo con la *EU/EC Directives: Commission Recommendation of 6 April 2005 on broadband electronic communications through powerlines 2005/292/EC.* Il gruppo di lavoro PLT studia i requisiti

	tecnici per evitare interferenza tra i vari utilizzatori dello spettro radio.
UPA	Universal Power Line Association ha proposto diversi standard per i sistemi PLC. Questi riguardano aspetti della tecnologia dei sistemi PLC. La prima area di intervento è la coesistenza, le cui specifiche sono state definite nel 2005. La seconda area di intervento è l'accesso a reti BB-PLC, le cui specifiche sono state definite nel 2006. La terza area di intervento è rivolta i sistemi e alle soluzione interne alle abitazioni ed è datata 2006.

Tabella 2. Standard ed enti di standardizzazione per PLC

TALQ

Il consorzio TALQ [14], fondato da aziende leader nell'ambito *Smart Lighting* (SL), si occupa di sviluppare un protocollo globalmente accettato per permettere ad un *Central Management System* (CMS) di controllare e monitorare molteplici *Outdoor Lighting Networks* (OLNs) al fine di evitare problemi di interoperabilità e integrazione di sistemi di SL.

Le OLN sono già utilizzate in diversi paesi. Il sistema consiste in un computer/server di controllo centrale (CMS) e un sistema di reti collegate a punti luce sul territorio (OLN).
Alcuni produttori hanno sviluppato le proprie tecnologie proprietarie, mentre altri stanno utilizzando protocolli simili. È perciò richiesta interoperabilità tra sistemi e componenti diversi per consentire agli utenti di beneficiare di sistemi di differenti produttori.
La mancanza di una comunicazione standardizzata tra CMS e OLN porta ad una situazione dove una città/regione si può trovare ad avere sistemi che non sono interoperabili tra loro e quindi difficili da integrare, gestire e manutenere (Figura).
L'assenza di uno standard globale ostacola l'adattamento di questi sistemi. Il Consorzio TALQ mira pertanto a standardizzare l'interfaccia tra il CMS e le OLNs (Figura 9).
In questo modo sarà possibile collegare OLNs di differenti tecnologie o fornitori ad un CMS anche esso di produttore generico.
L'interfaccia TALQ rappresenta una specifica per lo scambio di informazioni, adatta per l'implementazione in vari sistemi di trasporto fisici. TALQ si concentra sul livello di applicazione del protocollo di interfaccia tralasciando di definire i livelli fisici e di rete. Questo approccio favorirà la concorrenza e contribuirà a far crescere il mercato a beneficio degli utenti finali.

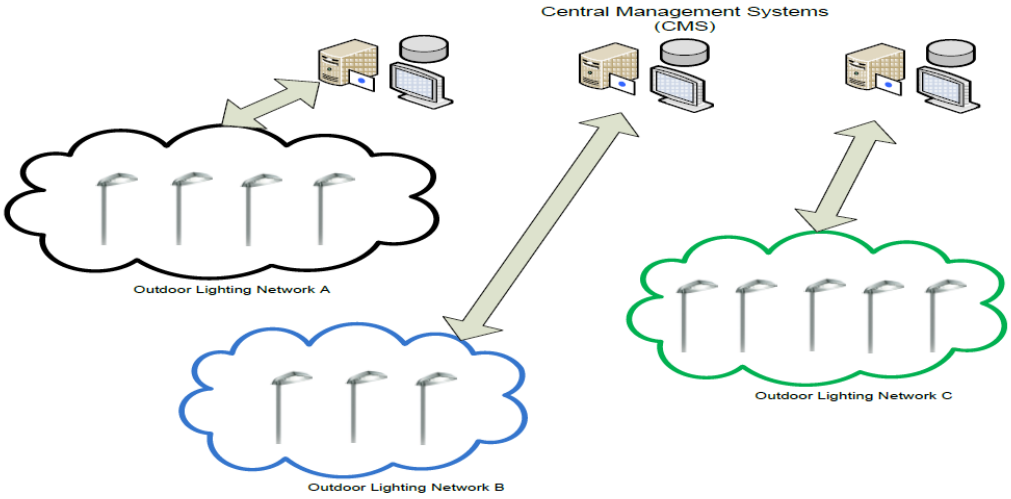


Figura 9 Architetture CMS e OLS

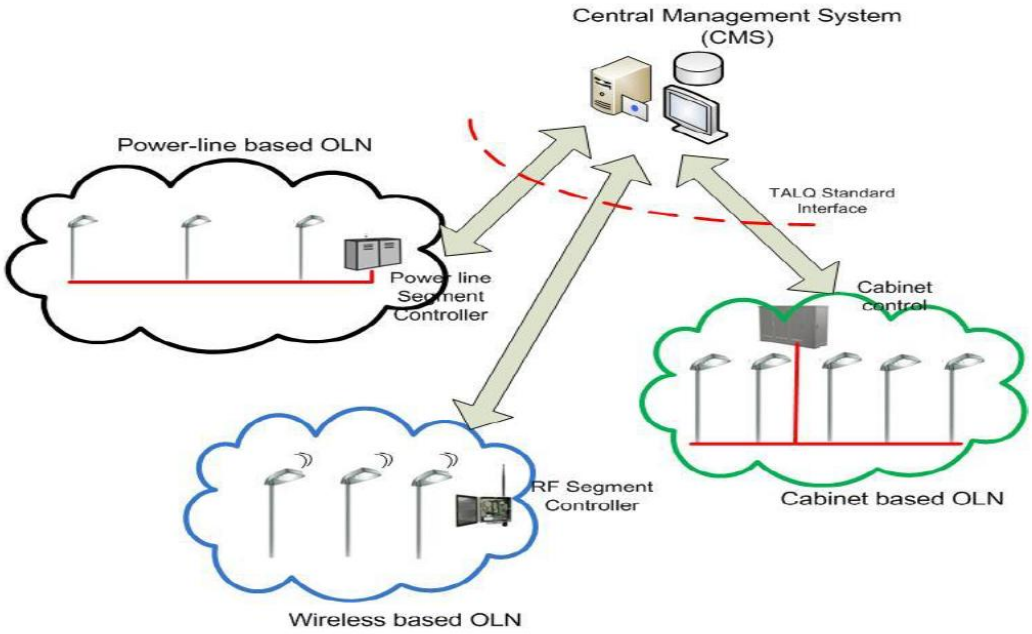


Figura 10 Interfaccia TALQ Standard con CMS e OLN

Dettagli maggiori su questo protocollo verranno discussi nella sezione 2.3 di questo documento.

Vantaggi e svantaggi nell'uso di una PLC

L'utilizzo delle PLC in uno scenario di Smart Cities presenta molti vantaggi, dovuti essenzialmente alla capillarità della rete elettrica e alla capacità di trasformarsi in infrastruttura di telecomunicazione. In [15] si fornisce una panoramica dei principali vantaggi, che possono essere qui riassunti:

- *Ampia copertura*: le linee di distribuzione elettrica possono essere usate per estendere la copertura richiesta dalle applicazioni di SG;
- *Basso costo*: le linee elettriche esistenti possono essere utilizzate per la comunicazione, senza ulteriori sforzi di cablaggio. Le PLC sono soluzioni efficienti per il monitoraggio da remoto dell'uso e del malfunzionamento della rete elettrica;
- *Flessibilità*: le reti PLC sono flessibili, forniscono un largo raggio di comunicazione;
- *Facilità di installazione*: l'installazione delle reti PLC è semplice per applicazioni indoor e possono essere usate per sistemi a basso costo;
- *Stabilità*: le PLC offrono comunicazioni stabili.

Nello stesso studio [15] si fa menzione anche di alcuni svantaggi dovuti al mezzo di trasmissione, che non è stato definito per la trasmissione di informazioni:

- *Presenza di fonti di rumore*: sulle reti di distribuzione elettrica sono presenti fonti di rumore quali motori elettrici, generatori e interferenze radio;
- *Capacità*: alcuni modem attualmente a disposizione hanno capacità limitate;
- *Mancanza di interoperabilità*: esistono diversi standard e protocolli che rendono difficile l'interoperabilità tra diverse PLC;
- *Presenza di attenuazioni e distorsioni dei segnali*: la topologia della rete e la fluttuazione dei carichi attenua e distorce il segnale sulle PLC;

- *Sicurezza*: nelle linee di distribuzione elettrica non sono utilizzati cavi intrecciati e schermati. Per questa ragione interferenze elettromagnetiche possono essere prodotte dalle PLC e dar luogo a problemi di sicurezza a livello di ricevitori di segnale.

2.2.3 Comunicazione a radiofrequenza

La comunicazione su canale wireless per il controllo di sistemi di SL è quella che viene utilizzata nei cyber physical systems. I protocolli più largamente utilizzati sono riportati in Tabella 3 e appartengono per lo più allo standard IEEE 802.1x. In particolare abbiamo il protocollo WiFi (IEEE 801.11) che consente ampia copertura, ma limitate capacità di organizzarsi in reti ad hoc. Più flessibile risultano i protocolli aderenti allo standard IEEE 801.15 (Bluestooth, ZigBee, Xbee, etc.) che consentono la creazione di reti mesh. Questo tipo di rete non prevede una struttura gerarchica centralizzata, ma tutti i nodi possono avere anche la funzione di relay e/o di router. Questa tipologia di reti risulta più dinamica della Wi-Fi in quanto si auto-organizza e auto-configura in maniera automatica al fine di stabilire connessioni a maglia stabili.

Protocollo	Standards	Max throughput	Banda	Max range
Wi-Fi	IEEE 802.11	54 Mbps	2.4/5 GHz	150 m
ZigBee	IEEE 802.15.4	20/40/250 Kbps	868/915 MHz 2.4 GHz	300 m
Bluetooth	IEEE 802.15.1	1/24 Mbps	2.5/5 GHz	150 m
RFID	ISO/IEC 2479	5/640 kbps	125 kHz.1356 MHz/ 433 Hz	10 cm-100 m

Tabella 3 Protocolli wireless comunemente utilizzati per il controllo di sistemi distribuiti

2.2.4 Vulnerabilità dei sistemi PLC

Criticità sulla sicurezza [16], [17]

Lo scenario delle Smart Grid (SG) è progettato per integrare dispositivi ad alto contenuto tecnologico capaci di gestire comunicazioni bidirezionali ad alta velocità per poter fornire servizi di gestione e monitoraggio. In tale contesto si inserisce anche lo Smart Lighting (SL).
Un tale sistema critico richiede l'implementazione di sistemi di protezione accuratamente definiti, capaci di rilevare tutte le possibili minacce e vulnerabilità. Tuttavia, le SG sono vulnerabili a numerosi rischi per la sicurezza in quanto per definizione, sono sistemi interconnessi atti a integrare innumerevoli dispositivi differenti. Di conseguenza, la consapevolezza dei requisiti di sicurezza e affidabilità si rende più evidente.
Esistono numerosi studi che vengono condotti per determinare i problemi di sicurezza e per migliorare le tecniche di difesa cyber delle SG in termini di requisiti e metodi di precauzioni basati su algoritmi e protocolli. Gli obiettivi di sicurezza delle SG riguardano la Disponibilità del sistema, l'Integrità dei dati e la Privacy. Analizzeremo nel dettaglio questi obiettivi, considerando che essi richiedono meccanismi di identificazione, autenticazione e controllo degli accessi sui sistemi interconnessi delle SG.

Disponibilità del sistema

La disponibilità del sistema si riferisce al garantire che persone o sistemi non autorizzati non possano negare l'accesso o l'utilizzo ad utenti autorizzati.
Per le SG, ciò si riferisce a tutti gli elementi di Information Technology (IT) presenti nell'impianto, come i sistemi di controllo, i sistemi di sicurezza, le postazioni di lavoro degli operatori, le

postazioni di lavoro degli ingegneri, i sistemi di esecuzione della produzione, così come i sistemi di comunicazione tra l'impianto ed il mondo esterno.
Gli attacchi malevoli relativi alla disponibilità del sistema possono essere considerati come attacchi Denial-of-Service (DoS). Questi tentano di ritardare, bloccare o corrompere la trasmissione dei dati attraverso i mezzi di comunicazione al fine di rendere le risorse di rete non disponibili. Una volta isolati i nodi sulla rete, non è possibile controllarli e monitorarli.
Dal momento che è ampiamente previsto che almeno una parte delle comunicazioni sulle SG sia basata su protocollo TCP/IP, questo rende effettiva la vulnerabilità agli attacchi DoS. Gli attacchi DoS contro il protocollo TCP/IP sono stati largamente studiati in letteratura riguardo le tipologie di attacco, le tecniche di prevenzione e le contromisure.
Tuttavia, una grande differenza tra una rete di comunicazione per SG e Internet è che per la prima il ritardo dei dati è più critico rispetto alla capacità di trasmissione dei dati, a causa dei vincoli temporali imposti per la gestione delle operazioni di controllo e monitoraggio. Per questo, il traffico di rete delle SG si presenta come *time-critical*.
Gli intrusi che hanno la possibilità di collegarsi ai canali di comunicazione, possono in maniera semplice effettuare attacchi di tipo DoS contro i nodi di rete delle SG. In particolar modo, sono molto sensibili a questo tipo di attacchi le comunicazioni wireless (attacchi *jamming*).
In definitiva, risulta di fondamentale importanza poter valutare l'impatto degli attacchi DoS sulla rete SG e di progettare contromisure efficaci a tali attacchi per provvedere alla continua disponibilità del sistema.

Integrità dei dati

Proteggere l'integrità dei dati si riferisce al prevenire la modifica dei dati scambiati dai nodi sulla rete della SG da parte di sistemi o persone non autorizzate.
Per i sistemi di comunicazione delle SG, questo concetto si applica alle informazioni relative alle letture dei sensori o ai comandi di controllo ed attuazione.
La violazione dell'integrità dei dati può causare problemi di sicurezza intesi come danneggiamento di attrezzature e macchinari, i quali possono essere pericolosi per la salute delle persone.
A differenza degli attacchi che mirano alla disponibilità delle risorse di rete, gli attacchi rivolti all'integrità dei dati possono essere considerati più sofisticati.
L'obiettivo degli attacchi sull'integrità dei dati può riguardare sia le informazioni personali degli utenti che utilizzano la rete, sia le informazioni sul funzionamento della rete stessa, ad esempio letture dei sensori, dati di videosorveglianza, etc.
In altre parole, tali attacchi tentano di modificare deliberatamente le informazioni originali scambiate dai nodi connessi nel sistema di comunicazione delle SG al fine di corromperli.
Il rischio di attacchi contro l'integrità dei dati è reale sia nello scenario SG che nel più specifico scenario di SL.
Un esempio notevole è il lavoro [18], che ha proposto un tipo di attacco, *false data injection*, capace di corrompere la stima dello stato in sistemi di sicurezza presenti nelle SG. In questo tipo di attacchi, si assume che l'attaccante abbia già compromesso i dati dei componenti connessi in rete e iniettato azioni di controllo malevoli al fine di provocare danno al sistema SG.

Privacy

Nelle SG, problemi di privacy riguardano le informazioni di consumo di energia derivate dall'attaccante e riguardanti gli utenti finali del sistema. I dati di consumo contengono

informazioni dettagliate sul comportamento di un cliente. Le comunicazioni su SG possono avere conseguenze impreviste per la privacy del cliente.
I dati sull'utilizzo di energia elettrica sono gestiti dagli *smart meters,* tali strumenti agiscono come canali laterali ricchi di informazioni, esponendo abitudini e comportamenti degli utenti. Alcune attività, come ad esempio guardare la televisione, hanno firme rilevabili di consumo di energia elettrica.
La storia ha dimostrato che, laddove gli incentivi finanziari o politici si allineano, le tecniche di estrazione dei dati sensibili si evolveranno rapidamente per soddisfare i desideri di coloro che vorrebbero sfruttare le informazioni [19].

Requisiti di sicurezza

L'affidabilità di una SG, così come di un sistema SL, dipende dai sistemi di controllo e comunicazione. Con il passare degli anni, l'evoluzione di questi sistemi ha permesso di gestire in maniera più sofisticata le dinamiche di comunicazione, consentendo un maggior controllo ed una maggiore affidabilità.
D'altro canto, più è alto il grado di connettività tra i nodi di un sistema di comunicazione, più deve essere sofisticato il sistema di sicurezza implementato per evitare violazioni.
In questa sezione verranno analizzati i requisiti di sicurezza informatica e fisica per gli scenari in esame.

Requisiti di sicurezza informatica

I sistemi di controllo industriale, come ad esempio i sistemi SCADA (*Supervisory Control And Data Acquisition*) vengono ampiamente utilizzati in scenari di SG. Diverse tipologie di attacchi informatici che mirano a perturbare il normale funzionamento dei sistemi sono avvenuti negli ultimi anni.
L'Idaho National Laboratory ha eseguito uno studio sperimentale nel 2007 per distruggere il meccanismo di controllo di un generatore diesel che veniva azionato a distanza.
In un altro episodio, si ritiene che l'esercito Russo abbia distrutto le utilità di sistema di SG georgiane tramite attacchi cyber durante il conflitto Russo-Georgiano del 2008.
Di sicuro, l'attacco informatico più significativo degli ultimi anni riguarda il famoso Stuxnet. Esso aveva come obiettivo i controllori a logica programmabile della Siemens, largamente utilizzati nei sistemi di controllo e comunicazione industriale. Il sistema operativo del controllore era in esecuzione su Windows e Stuxnet sfruttava quattro diverse vulnerabilità zero-day e come risultato causò la compromissione del 60% degli impianti di energia nucleari Iraniani [20].
Questi attacchi dimostrano che la sicurezza informatica della SG svolge un ruolo fondamentale e dovrebbe essere presa in considerazione in modo accurato.
I requisiti di sicurezza informatica sono in primo luogo correlati alla conoscenza della minaccia.
Di conseguenza, gli attacchi devono essere rilevati e contromisure immediate devono essere eseguite contro l'attacco.
Tuttavia, poiché la SG è concepita per essere distribuita su ampie zone con numerosi nodi connessi è da considerarsi come una rete di comunicazione aperta che rende impossibile garantire la sicurezza di ogni nodo contro attacchi.
Per la rete di comunicazione risulta necessario poter verificare e confrontare i cambiamenti nel flusso di dati, monitorando e rilevando eventuali anomalie nel traffico di rete.

Ad esempio, gli attacchi DoS mirano a saturare le risorse di trasmissione dati inviando illimitate false richieste verso i nodi del sistema.
Eventuali attacchi agli *smart meters* possono causare la perdita di dati sui consumi e possono causare gravi ripercussioni economiche.
Pertanto, l'integrità delle informazioni di fatturazione, dati dei contatori, comandi di gestione, sono molto importanti.

Requisiti di sicurezza fisica

Le più importanti apparecchiature fisiche presenti nelle SG sono gli *Smart Meters* (SM) e le *Phasor Measurement Unit* (PMU) che rappresentano nodi di comunicazione sulla rete.
Gli SM che sono misuratori di consumo elettrico e sono gestiti con tecnologie di comunicazione bidirezionale per trasmettere i tassi di consumo e dati sui prezzi in tempo reale.
Le PMU sono destinate a misurare la qualità di potenza di qualsiasi sistema connesso, e le tecnologie di comunicazione sono generalmente basate su apparati wireless.
Il *Metering Data Management System* (MDMS), ovvero il sistema di gestione dei dati di misurazione, comunica con i sensori ed i centri di controllo attraverso la misura ed il controllo delle reti al fine di analizzare i dati attuali e prendere decisioni.
Data la criticità di questi sistemi, un eventuale guasto/attacco fisico potrebbe compromettere il normale funzionamento dell'intera SG. Per tale motivo è necessario uno studio accurato dei metodi di prevenzione per rendere sicuri e robusti i sistemi interconnessi sulle SG. È altresì necessario poter valutare le potenziali conseguenze a cascata di danni fisici ai sistemi, intervenendo con appropriate contromisure.

Soluzioni attuali e contromisure [21]

In questa sezione vengono esaminate diverse soluzioni esistenti in materia di sicurezza informatica per le comunicazioni sulle SG. Tale analisi verrà contestualizzata considerati gli obiettivi di sicurezza descritti precedentemente, ovvero la disponibilità del sistema, l'integrità dei dati e la privacy.

Disponibilità del sistema

A causa dell'interazione delle reti di informazione e dei dispositivi elettrici sui sistemi energetici, la SG deve essere in grado rilevare e contrastare gli attacchi DoS che possono essere lanciati ovunque nelle reti di comunicazione. Il rilevamento degli attacchi è il primo passo verso l'attuazione di contromisure contro questi attacchi.
La maggior parte dei metodi di rilevamento di attacchi DoS appartengono alla rilevazione passiva che monitora continuamente lo stato della rete, come ad esempio il rapporto tra il carico di traffico e la trasmissione di pacchetti: un allarme viene attivato una volta che vi è una discrepanza evidente tra i nuovi campioni ed i dati storici.
La metodologia esistente per il rilevamento di attacco DoS può essere applicata direttamente alle reti di comunicazione nelle SG.
Gli approcci più utilizzati per mitigare gli attacchi DoS sono progettati per il livello di rete e molti di loro hanno dimostrato di essere efficace per Internet. Ad esempio, i seguenti meccanismi sono discussi ampiamente [22]:

- Limitazione di banda: L'idea di base di meccanismi di limitazione di banda è di imporre un limite di velocità su un insieme di pacchetti che sono stati caratterizzati come potenzialmente dannosi dal meccanismo di rilevamento. Di solito è utilizzato quando il meccanismo di rilevamento ha molti falsi positivi o non è in grado di caratterizzare con precisione il flusso di attacco.
- Filtraggio: comprovato con metodi di rilevamento di attacco, meccanismi di filtraggio possono confrontare gli indirizzi sorgenti dei pacchetti con la lista nera fornita dai rilevatori di attacco per filtrare tutti i flussi sospetti. Di conseguenza, i pacchetti che provengono da potenziali attaccanti non saranno inoltrati ulteriormente o indirizzati alle vittime.
- Riconfigurazione: al fine di mitigare l'impatto degli attacchi DoS, una soluzione è quella di riconfigurare l'architettura di rete, come ad esempio cambiare la topologia di rete della vittima o di isolare le macchine sotto attacco. Tuttavia, potrebbe non essere facile usare meccanismi di riconfigurazione, dal momento che sezioni della rete SG sono statiche a causa della topologia fissa di distribuzione di energia e delle apparecchiature di trasmissione.

Integrità dei dati

Diversi modelli di politica di integrità dei dati (ad esempio, Biba [23], LOMAC [24], e Clark-Wilson [25]) sono stati sviluppati per governare i livelli di integrità di un sistema. Il modello di Biba assicura che i processi non possono danneggiare i dati a livelli più alti e non sono danneggiati dai dati di processi di livello inferiore [23]. Il modello LOMAC imposta dinamicamente il livello di integrità di un processo al livello minimo di integrità dei dati con cui interagisce [24]. Allo stesso modo, il modello Clark-Wilson permette un di scartare o aggiornare il livello di integrità dei dati permettendo così di interagire con i dati a livelli di integrità inferiore [25].
L'integrità dei dati dipende dalla genuinità del processo e dei dati raccolti. Nel complesso, l'integrità dei dati richiede una catena di fiducia. Garantire l'integrità dei dati generati richiede assicurare l'integrità del processo di generazione e l'integrità dei dati di ingresso al processo. Garantire l'integrità dei dati di input richiede garantire la genuinità del processo di comunicazione o il dispositivo di input.
La valutazione dell'integrità dei dati comporta la verifica della sorgente e richiede la conoscenza delle *fingerprints*: sequenze alfanumeriche o stringhe di bit di lunghezza prefissata utilizzate per identificare in maniera sicura le informazioni con le caratteristiche intrinseche delle stesse. La gestione dell'integrità di un sistema interattivo completo, come quello delle SG, è un compito impegnativo, in quanto possono essere richieste migliaia di misurazioni e la conoscenza delle loro *fingerprints*.

Privacy

La privacy delle misure è richiesta per garantire l'identificazione e il controllo di accesso al fine di evitare accessi non autorizzati ai dispositivi.
Le comunicazioni nelle SG devono assicurare che i dati di comunicazione preservino la privacy ovunque e in qualsiasi momento.
In [11], gli autori hanno descritto un metodo per rendere anonimi in modo sicuro frequenti (ad esempio, ogni pochi minuti) dati di misura elettrici inviati da uno *smart meter*. Sebbene tali dati frequenti di misurazione possano essere richiesti da un provider di energia elettrica per ragioni

operative, potrebbero non necessariamente essere attribuibili ad un consumatore specifico. Tuttavia, essi hanno bisogno di essere attribuibili in modo sicuro ad una posizione specifica (ad esempio, un gruppo di case o appartamenti) all'interno della rete di distribuzione elettrica. Questo metodo non preclude la fornitura dei dati di misura imputabili che risultano necessari per altri scopi quali la fatturazione, gestione del conto o marketing.
Le tecnologie *Public Key Infrastructure* (PKI) sono considerate per soddisfare la prevenzione e i requisiti di privacy. La PKI include diverse politiche e procedure per identificare i clienti in base alla certificazione digitale che definisce le strategie di gestione, configurazione e il funzionamento.
Un metodo per fornire la privacy è utilizzando le funzioni di crittografia su tutta la rete di comunicazione per crittografare e decrittografare i dati.
Le tecnologie *Public Key Infrastructure* (PKI) sono considerate per soddisfare la prevenzione e i requisiti di privacy. La PKI include diverse politiche e procedure per identificare i clienti in base alla certificazione digitale che definisce le strategie di gestione, configurazione e il funzionamento.
Il certificato PKI deve essere controllato e verificato ad ogni ciclo riferendosi alle richieste di firma. Anche se le tecnologie PKI sono piuttosto complesse, esistono implementazioni nelle SG che ne sfruttano le qualità di sicurezza.
Oltre ai meccanismi di PKI, la sicurezza può essere ottenuta utilizzando vari protocolli di comunicazione che consentono un flusso di dati efficiente ed affidabile sulla rete della SG.
In questo contesto, un'applicazione di esempio è il protocollo *Secure Smart-Metering* (SSMP) che viene eseguita in [27] ed è stata progettata per tecnologia PLC. Il protocollo si basa su diversi sistemi crittografici quali PKI, certificati digitali, chiavi di autenticazione e chiavi condivise. Questo protocollo è in grado di impedire accessi non autorizzati, attacchi DDoS, sostenendo la privacy su una rete PLC [27].
Poiché la tecnologia PLC non richiede una infrastruttura di comunicazione dedicata, può essere facilmente implementata per permettere la connettività tra nodi di una rete.
D'altra parte, è importante garantire prestazioni in tempo reale e operazioni della SG a seconda del metodo di comunicazione selezionato.
Alcuni algoritmi di comunicazione piuttosto complessi per le infrastrutture di trasmissione e distribuzione possono causare un abbassamento delle prestazioni e aumentare i problemi di vulnerabilità imprevisti [2].
Pertanto, il trade-off tra sicurezza e prestazioni deve essere sempre preso in considerazione.

2.2.5 Criticità dei sistemi di monitoraggio e controllo utilizzando internet

Gli strumenti di controllo, installati in ogni singolo lampione, vengono gestiti da remoto attraverso un Server *Cloud* sfruttando l'assegnazione di indirizzi IP univoci ad ogni componente di telecontrollo connessa [28]. Questo permette da un lato la possibilità di monitorare e configurare l'intero sistema di *Smart Lighting* da qualsiasi postazione con accesso sulla rete, dall'altro rende vulnerabili questi sistemi a possibili azioni malevoli attraverso internet.

Phishing

Il *phishing* rappresenta uno dei crimini più redditizi per attori malevoli su internet. Il termine "phishing" deriva dalla analogia del "fishing" (= atto del pescare) per le password e le credenziali delle vittime nel web. La "ph" deriva da "phone phreaking", che era una tecnica utilizzata per attaccare i sistemi telefonici negli anni '70. La parola "phishing" è stata utilizzata per la prima volta

su Internet da un gruppo di hacker nel 1996, che entrò in possesso di utenze dell'America Online (AOL) ingannando gli utenti e portandoli a rivelare le loro password [29].
Il *phishing* può essere considerato come un furto di identità automatizzato, che prende il vantaggio della natura umana e di Internet per ingannare milioni di persone.
Le analisi del gruppo Gartner hanno evidenziato nel mese di aprile 2004, che circa 1.8 milioni di americani avevano già fornito i propri dati ad un *phisher*. È stato osservato che negli ultimi anni gli attacchi di *phishing* sono cresciuti rapidamente e rappresentano una vera e propria minaccia per la sicurezza globale.
Le campagne di *phishing* cercano di estrarre i dati segreti dalle vittime e possono portare a notevoli perdite finanziarie. Studi hanno dimostrato che un terzo di tutti i tentativi di *phishing* nel 2013 sono stati destinati verso conti bancari o per ottenere altre informazioni finanziarie [29]. Dal 2012, gli attacchi di *phishing* finanziari sono aumentati dell'8,5% rispetto al 2011, un massimo storico [16].
L'obiettivo principale di queste campagne è quello di sfruttare le vulnerabilità presenti nei sistemi connessi in internet, di natura tecnica o a causa di inconsapevolezza degli utenti. Ciò significa che i ricercatori devono fornire strumenti di difesa contro questi attacchi, sia a livello tecnico che di consapevolezza dell'utente.
A dispetto del fatto di causare gravi danni finanziari agli utenti attraverso Internet, *spam* e *phishing* sono ancora in crescita ad un ritmo veloce e continueranno a farlo fino a quando 1 su 100.000 destinatari risponde effettivamente alle frasi come *"Clicca qui"* nelle e-mail di spam.
Secondo i report dell'Anti-Phishing Working Group (APWG), il *phishing* continuerà a crescere con l'utilizzo delle tecnologie più avanzate e diventerà la principale minaccia su Internet, superando lo *spam* [17].

Motivazione

I *phisher* sfruttano sempre la natura umana che generalmente ignora i messaggi di avviso importanti. La mancanza di consapevolezza circa gli attacchi di *phishing* nella società è anche il motivo principale per cui gli attacchi di *phishing* hanno incontrato così tanto successo.
Ogni volta che qualsiasi ricercatore sviluppa una certa tecnica per prevenire questi attacchi, i *phisher* cercano di scoprire una nuova scappatoia per effettuare attacchi con successo.
Ricordando il fatto che il *phishing* viene utilizzato principalmente per i guadagni finanziari, ci sono altri fattori che motivano il *phisher* a commettere il reato. Le motivazioni che stanno dietro queste attività sono le seguenti:

- Il furto di credenziali di accesso: il *phisher* ruba le credenziali di accesso di servizi on-line come eBay, Amazon e Gmail utilizzando la posta elettronica come esca con un messaggio di avviso di cambiare la password e fornendo un collegamento ipertestuale.
- Furto di credenziali bancarie: le credenziali di accesso on-line e della carta di credito come il numero di carta, date di scadenza ed emissione, il nome del titolare, il numero CCV.
- Acquisizione dei dati personali: i dati personali, come l'indirizzo e il numero di telefono, sono altamente vendibili e in costante richiesta da parte delle imprese di marketing diretto.
- Furto di segreti commerciali e documenti riservati: i *phisher* stanno prendendo di mira organizzazioni specifiche per l'acquisizione di informazioni proprietarie per utilizzarle direttamente o per venderle alle parti interessate.

- Fama e notorietà: un aspetto psicologico molto interessante del *phishing* è quando l'informazione oggetto di *phishing* non risulta a scopo di lucro, ma principalmente per ottenere riconoscimento e notorietà tra i propri contatti.
- Propagazione di un attacco: attraverso il *phishing* gli agenti malevoli possono utilizzare un singolo host come un nodo interno compromesso per un attacco futuro.

Ciclo di vita del phishing

Le seguenti fasi sono successive in un attacco di *phishing*:
Fase 1: Progettazione e configurazione
Nella prima fase, gli aggressori identificano l'organizzazione di destinazione, l'individuo o una nazione. Successivamente, il loro compito è quello di ottenere i dettagli circa l'organizzazione e la rete. Questo può essere effettuato visitando il luogo fisicamente o monitorando il traffico in entrata e in uscita dalla rete. Il passo successivo è quello di impostare gli attacchi utilizzando un mezzo possibile, ad esempio, un sito web o una e-mail con link malevoli, che possono reindirizzare la vittima a qualche pagina web di frode.
Fase 2: Phishing
Il passo successivo è quello di inviare e-mail false, per esempio, mascherate da qualche organizzazione bancaria nota alla vittima, utilizzando gli indirizzi e-mail raccolti e che richiedono all'utente di aggiornare alcune informazioni con urgenza cliccando su qualche link maligno. Le e-mail possono essere inviate a più persone o a una persona specifica di un'organizzazione.
Fase 3: Break-in / infiltrazione
Non appena la vittima apre il collegamento di frode, un malware viene installato sul sistema che permette al malintenzionato di intromettersi e cambiare i propri diritti di configurazione. In altri casi, potrebbe portare a qualche falsa pagina che richiede le credenziali.
Fase 4: La raccolta dei dati
Una volta che gli aggressori hanno accesso al sistema dell'utente, i dati richiesti vengono estratti, e questo può portare a perdite finanziarie per la vittima. In caso di attacchi malware, l'attaccante è adesso in grado di ottenere l'accesso remoto al sistema.
Fase 5: Rimozione prove
A questo punto, dopo aver ottenuto le informazioni richieste, il *phisher* rimuove tutte le prove delle sue azioni malevoli.

Tassonomia di attacchi di phishing

Gli attacchi di *phishing* sono classificati in due categorie: attacchi di ingegneria sociale e attacchi di *phishing* basati su malware. Il *phishing* basato su ingegneria sociale tenta di acquisire le credenziali delle vittime utilizzando qualche falso sito web o l'invio di messaggi di posta elettronica fasulli che sembrano essere legittimi per ingannare l'utente [33]. Allo stesso modo il *phishing* basato su malware utilizza una varietà di programmi maligni, che sono principalmente software indesiderato in esecuzione sul sistema di destinazione.

Phishing basato sull'ingegneria sociale

Gli attacchi di *phishing* basati sull'ingegneria sociale intendono acquisire identità o altre informazioni riservate della vittima tramite e-mail contraffatte o false. Gli attacchi di ingegneria

sociale sono portati in azione con movente simile a quello dell'*hacking*, cioè, con lo scopo di acquisire l'accesso illegale ad un sistema o ottenere informazioni riservate su un'organizzazione o un individuo. Gli obiettivi comuni sono le grandi aziende, i militari e le agenzie di governo [34].
I bersagli degli attacchi di ingegneria sociale agiscono a due livelli: fisico e psicologico [35]. Il primo riguarda il luogo fisico dove effettuare l'attacco e possono interessare il luogo di lavoro, il telefono ed il web. Gli agenti malevoli focalizzano la ricerca di un modo per creare un ambiente psicologico favorevole con la vittima per far sì che l'attacco abbia successo. Nonostante il metodo utilizzato, l'obiettivo primario è quello di far credere alla vittima che l'attaccante sia una persona fidata in modo che gli possa fornire le informazioni richieste [36]. I *phisher* che sfruttano questa tecnica non puntano mai ad ottenere molteplici informazioni in una sola volta da un utente. Piuttosto, essi cercano di ottenere piccoli dettagli da molte persone, al fine di guadagnare la loro fiducia.
Le tecniche di ingegneria sociale utilizzate per scopi malevoli sono le seguenti:

- *Phishing* attraverso siti web: è un attacco il cui obiettivo è quello di colpire una persona in particolare, piuttosto che un sistema. Questi attacchi sono molto facili da mettere in atto a causa del fatto che la creazione di un sito di *phishing*, come replica esatta di un certo sito legittimo, non è un problema per l'attaccante. L'obiettivo principale è quello di truffare le persone al fine di ottenere i propri dati personali e finanziari.
- E-mail *phishing*: Il primo passo del *phisher* è quello di creare un sito di *phishing*; successivamente vengono inviati numerosi messaggi di posta elettronica falsi che possono chiedere all'utente di cliccare su un link all'interno di un messaggio e possono dare via al furto di identità o altre informazioni personali. Il *phisher* utilizza quindi l'identità della vittima per ottenere benefici finanziari illeciti o per qualche altro scopo. Nonostante il fatto che le email di *phishing* siano state sviluppati in uno stile più riconoscibile nel corso del tempo, ci sono ancora una serie di misure o indizi che indicano la loro natura ingannevole.
- *Spear Phishing* [37]: è una delle più grandi minacce informatiche di ingegneria sociale. Il *phisher* maschera l'e-mail come se inviata da qualcuno nella stessa organizzazione o un conoscente della vittima. Tale stratagemma aumenta la probabilità di successo per l'attacco poiché le vittime risultano più disponibili ad aprire un'e-mail se sembra provenire da qualcuno di familiare. Ciò rende questi attacchi molto difficili da individuare e l'attaccante può avere successo nella violazione dell'organizzazione senza che nessuno lo sappia. Gli attacchi di *spear phishing* iniziano con l'attaccante che raccoglie informazioni riguardante l'organizzazione e finiscono con l'ottenere accesso alle informazioni richieste.

Phishing basato su malware

Il malware è un software dannoso che di solito è installato su una macchina senza la consapevolezza della vittima e talvolta anche la vittima può essere indotta a scaricare software anti-virus, mentre in realtà rappresentano virus o malware. Questo software dannoso può accedere ai dati confidenziali degli utenti e inviarli al *phisher*. Il malware sfrutta le lacune dei sistemi software browser e di funzionamento, o fa uso di tecniche ingannevoli per incoraggiare la vittima ad eseguire codice malevolo. Secondo APWG, nel primo trimestre del 2013 [38], circa 5 milioni di tipologie di malware sono stati segnalate dalla società PandaLabs che ha aumentato la conta di nuovi campioni di malware a 15 milioni. I Trojan restano il malware più comunemente usato, che è il 72% di tutti i campioni di malware trovati. Il numero totale di sistemi infetti in tutto il mondo è di circa 33%. La Cina ha il più alto numero di sistemi infetti.

Tassonomia delle tecniche di difesa contro il phishing

Le email di *phishing* rappresentano una sorta di *spam*, un meccanismo criminale basato su richieste di posta elettronica false. L'obiettivo fondamentale è quello di rubare le informazioni personali o riservate dalle vittime. In questa sezione, si discutono approcci per filtrare le email di *phishing* e di rilevare le pagine web malevoli.

Educazione dell'utente

Educare l'utente si riferisce a diffondere la consapevolezza e la comprensione sul *phishing* tra gli utenti Internet. Approcci di formazione offrono informazioni on-line sui rischi di tali attacchi e le loro tecniche di prevenzione [38]. Alcuni approcci forniscono anche formazione on-line e test per gli utenti.
Downs et al. [39] hanno eseguito uno studio sulla risposta degli utenti agli attacchi *phishing* e hanno concluso che l'utente deve essere istruito sul *phishing*, piuttosto che essere messo in guardia sulle conseguenze negative degli attacchi. In questo studio, 232 utenti di computer sono stati invitati a visualizzare alcuni messaggi di posta elettronica e rispondere a qualche domanda relativa ad essi. Questo studio ha dimostrato che l'utente che possiede conoscenze sugli URL ha meno probabilità di cadere in attacchi di *phishing*, mentre la comprensione di altri strumenti web, ad esempio, i cookie e software dannosi, non ha ridotto il rischio di cadere in questi attacchi.
Huang et al. [40] hanno studiato che gli utenti diventano vittime di questi attacchi sia perché non sono in grado di distinguere tra un sito web legittimo e uno di *phishing*, sia per ignoranza delle avvertenze e gli indicatori delle barre degli strumenti.
Shen et al. [41] hanno mostrato alcune delle caratteristiche indirette, vale a dire, gli utenti di sesso femminile hanno più probabilità di cadere vittime di attacchi di *phishing* rispetto ai maschi. Lo stesso vale per le persone tra i 18 ei 25 anni di età a causa della mancanza di consapevolezza e di conoscenza tecnica.
Dong et al. [42] hanno proposto un modello che descrive l'interazione dell'utente riguardo al processo decisionale, che inizia non appena l'utente visualizza il messaggio di *phishing* o una pagina web e si ferma quando l'utente termina la sua attività. L'obiettivo è quello di individuare le campagne di *phishing* attraverso la comprensione del modo in cui gli utenti reagiscono al *phishing* sulle pagine web o sulla posta elettronica.
La formazione degli utenti di Internet per identificare i messaggi di posta elettronica e siti web dannosi può essere molto efficace nel prevenire attacchi di *phishing*.
Negli ultimi anni, sono stati proposti vari metodi per formare gli utenti online, test e giochi di formazione si sono dimostrati come un metodo efficace.
Kumarguru et al. [43] hanno proposto un metodo in cui le comunicazioni formative sono inviate agli utenti a intervalli fissi. Essi propongono anche un metodo di formazione integrata nelle attività quotidiane degli utenti in modo che essi non siano tenuti a leggere da altre fonti esterne.

Metodi di difesa basati sul software

Le principali tecniche di difesa contro il *phishing* basate su software sono le seguenti:
Protezione a livello di rete: in questo approccio, ad un certo intervallo di indirizzi IP o ad un insieme di domini non è permesso entrare nella rete. In [30], i DNSBL fanno uso del protocollo DNS e aggiornamenti vengono creati regolarmente osservando il traffico di rete. Un *Intrusion Detection*

System (IDS) open-source come *Snort* può essere utilizzato a livello di rete per verificare il traffico di pacchetti.
Meccanismi basati su autenticazione: in questo approccio, si conferma se il messaggio è stato inviato da un determinato percorso e se il nome di dominio risulta valido. Queste tecniche aumentano la sicurezza delle comunicazioni e-mail. Gli schemi di autenticazione sono abbastanza semplici e possono essere eseguiti sia a livello di dominio che applicando la firma digitale al documento prima di inviarlo. D'altro canto, questi metodi richiedono che la stessa tecnologia sia utilizzata dal mittente e dal destinatario. Un altro meccanismo di autenticazione chiamato *transaction authentication numbers* viene utilizzato dalle banche. Quest'ultimo però non garantisce la sicurezza da attacchi di tipo *Man-In-The-Middle* (MITM) ed è costoso in termini di tempo e di calcolo [45].
Strumenti lato Client: questi includono il filtro sul profilo utente e le barre degli strumenti dei browser. Essi dipendono anche da blacklist e whitelist, tecniche in cui un elenco di *phishing* riconosciuti e di siti legittimi viene scaricato con aggiornamenti a intervalli regolari. I limiti di queste tecniche sono che non riescono a rilevare attacchi zero-day.
Blacklist: liste che contengono URL e indirizzi IP sospetti e vengono frequentemente aggiornate. Esse non forniscono alcuna protezione da attacchi di *phishing* zero-day e possono rilevare solo il 20% di questi attacchi. Gli studi condotti concludono che 47-83% di URL di *phishing* sono nella lista nera dopo 12 h. Questo ritardo è significativo in quanto il 63% degli attacchi di *phishing* termina entro le prime 2 ore [46].
Filtri lato server e classificatori: questi sono basati su approcci di filtraggio dei contenuti e sono appropriati per combattere attacchi zero-day. Questi filtri sono basati su tecniche di apprendimento automatico e data mining.

2.2.6 Il Sistema Smart Street in ENEA

Figura 11. Area interessata dal progetto Smart Street ENEA (Fonte UVAX [47])

Il progetto Smart Street di ENEA realizza un sistema di strada intelligente all'interno dello Smart Village nella sede di ENEA – Casaccia.
Esso si compone di una strada illuminata e sensorizzata in maniera intelligente, come è possibile vedere in Figura 11.
Il sistema è costituito da una serie di venti lampione con armature LED Ampera Midi Schreder, equipaggiate con nodi di telecontrollo BPLC UVAX, in particolare la realizzazione di esso ha previsto l'installazione di

- Telecontrollo punto-punto a larga banda ad onde convogliate
  - 2 quadri
  - 20 nodi cablati all'interno degli Ampera

    - 1 nodo di comando in morsettiera per una linea di 4 globi a parete
    - 22 nodi cablati a palo per apparecchi esistenti
    - safd
- 2 telecamere Smart Eye per l'analisi scena
- 1 telecamera Smart Eye per la sicurezza
- 1 Hot spot WiFi.

L'obiettivo dell'installazione è il test di servizi a banda larga [47]:

- Telecontrollo dell'illuminazione illuminazione punto-punto con regolazione del flusso in tempo reale;
- Videosorveglianza analisi video
- Illuminazione adattiva (traffico, meteo, emergenza)
- Servizio WiFi

I dispositivi utilizzati sono i BPLC UVAX.

2.2.7 Architettura del sistema di controllo remoto e dispositivi

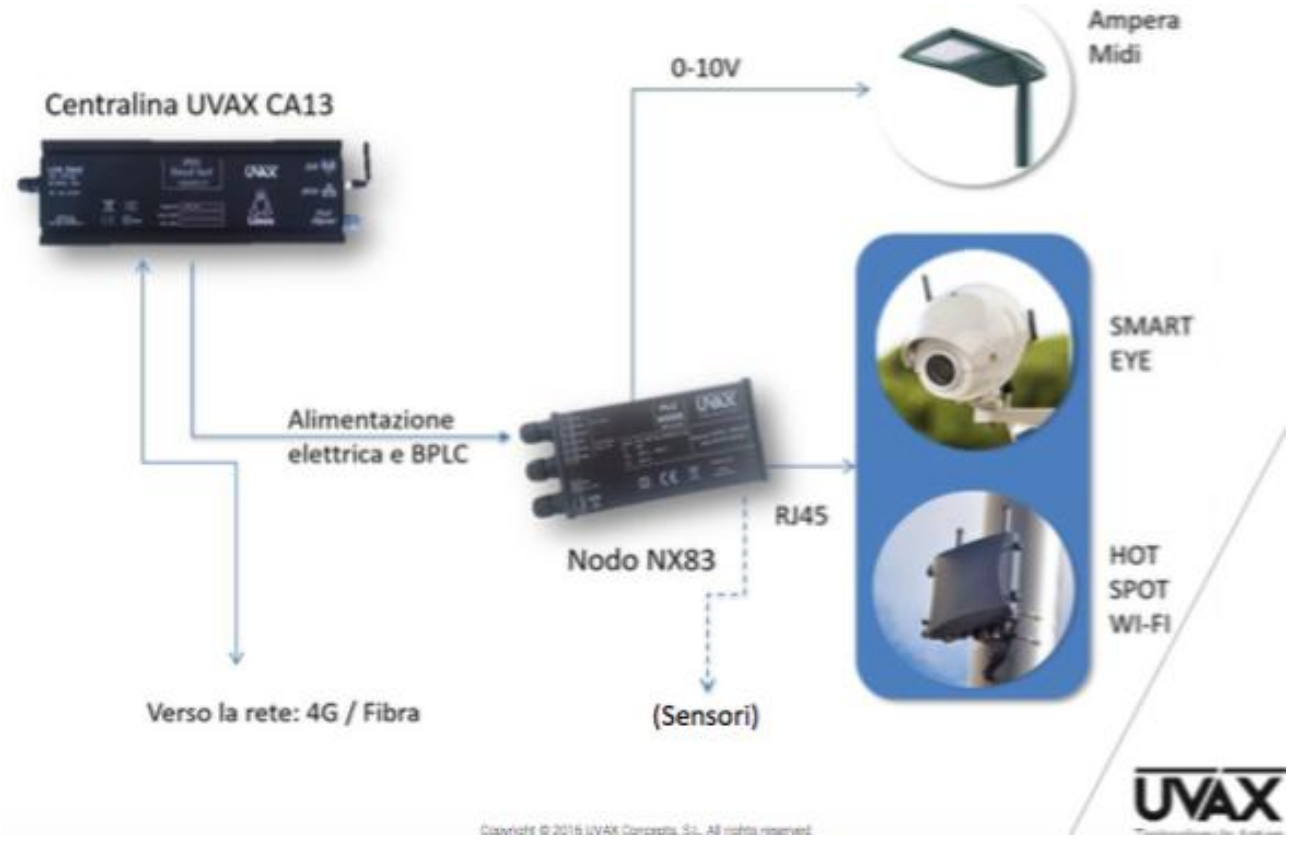


Figura 12 Architettura del sistema di controllo remote (Fonte UVAX [47])

L'architettura per implementare il sistema Smart Street si compone essenzialmente di tre elementi (si veda Figura 12):

- Nodi
- Concentratori
- Central Management Software (CMS).

Da un punto di vista software, dopo l'installazione, il gestore del sistema accede all'infrastruttura tramite software di rete in maniera sicura mediante autenticazione e cifratura e può impostare, configurare ogni singolo elemento della rete.

Dal punto di vista hardware il sistema di compone di nodi e concentratori che giocano ruoli diversi all'interno della rete.

Concentratori

I concentratori (Figura 13) sono installati all'interno delle cabine elettriche (2 nel caso del progetto Smart Street). Essi controllano i nodi direttamente collegati tramite rete ad onde convogliate e si interfacciano con il backbone della rete del management center.

Figura 13 Nodo concentratore (Fonte UVAX [48])

Le specifiche tecniche e le certificazioni si trovano in Tabella e Tabella rispettivamente.

Electrical Specifications		
Input		
	Input Voltage Range	100...277 VAC
	Input Frequency	50...60 Hz
	Power Factor	> 0.80
Coupling Output	Maximum Power	15 W
	Maximum Output Voltage	4,5 V(RMS)
	Maximum Output Current	70 mA
Network Link		
	LAN	Ethernet
	Wiring	CAT-5

Environment		
Use		
	Indoor	IEC 60529, IP43
	NEMA3	Type 1
Temperature	Operating / Ambient	-25 °C.....60 °C
	Storage	-25 °C.....85 °C
	Tc (Fig1)	65 °C

Tabella 4 Specifiche tecniche dei concentratori (Fonte UVAX [48])

2006/95/CE	EN60950-1: 2006+A11: 2009 EN60529_ 1991+A1: 2000
2004/108/CE	EN55022:2006+A1:2007 EN61000-3-2:2006 EN61000-3-3:2008 EN61547: 1995+A1: 2000 TGN17
UL	UL 916, 4 Edition, 2010-06-04
FCC	FCC CFR47 PART 15 SUBPART B ICES-003 ISSUE 5

Tabella 5 Certificazioni dei concentratori (Fonte UVAX [48])

Nodi

I nodi (Figura 14) sono installati sui singoli corpi illuminanti e sono collegati tramite rete ad onde convogliate ai concentratori. Essi sono dotati di indirizzo IP unico e si comportano da modem PLC, inviando e ricevendo informazioni dalla powerline, senza ulteriori necessità di cablaggio.
I nodi controllano i corpi illuminanti, consentendo la variazione dell'illuminazione, ma possono integrare dispositivi diversi come:

- GPS
- Ripetitori Wi-Fi
- Contatori intelligenti
- Sensori di temperatura e luce
- Telecamere.

Figura 14 Nodo (Fonte UVAX [48])

Le specifiche tecniche e le certificazioni si trovano in Tabella e Tabella rispettivamente.

Electrical Specifications				Environment		
Input				***Use***		
	Input Voltage Range	100...277 VAC			Indoor	IEC 60529, IP65
	Input Frequency	50...60 Hz			NEMA3	Type 1
	Power Factor	> 0.80				
	Maximum Power	4 W		***Temperature***		
					Operating / Ambient	-25 ºC.....60 ºC
Output VAC	Output Voltage range (Vac)	100..277 VAC			Storage	-25 ºC.....80 ºC
	Maximum output current	4A			Tc (Fig2)	65 ºC
	Maximum output power	400W				

Tabella 5 Specifiche tecniche dei nodi (Fonte UVAX [48])

2006/95/CE	EN60950-1: 2006+A11: 2009
	EN60529_ 1991+A1: 2000
2004/108/CE	EN55022:2006+A1:2007
	EN61000-3-2:2006
	EN61000-3-3:2008
	EN61547: 1995+A1: 2000
	TGN17
UL	UL 916
FCC	PART 15 / PLC

Tabella 6 Certificazioni dei nodi (Fonte UVAX [48])

I nodi ed i concentratori hanno lo stesso firmware e implementano lo stesso livello fisico. Per trasmettere utilizzano entrambi la OFDM, che come si è già avuto modo di notare, presenta numerosi vantaggi in termini di robustezza dovuti all'approccio a bassa frequenza della modulazione. Questo infatti consente da una parte di semplificare l'equalizzazione del canale, dall'altra di trasmettere dati attraverso cavi in rame caratterizzati da interferenze e multipath. In particolare viene utilizzato la modalità High Ultra Reliable Transmission OFDM (HURTO)[49], in cui i canali vengono selezionati in maniera adattativa rispetto alla qualità e le trasmissioni ridondate in modo da garantire la qualità del servizio.

2.2.8 Topologia della rete

Il sistema Smart Street installato nello Smart Village ENEA – Casaccia ha la capacità di auto-configurarsi con una topologia ad albero, la cui radice è rappresentata dai concentratori. Ogni nodo, infatti, ha un numero limitato di porte di comunicazione PLC, pertanto l'organizzazione ad albero consente di massimizzare la connessione tra un numero alto di nodi nella rete. I livelli degli alberi vengono aumentati quando un nodo *parent* ha esaurito la sua capacità di connessione: un nuovo nodo che appare sulla rete sceglie come riferimento una foglia cui connettersi.
In Figura 15, ad esempio, il nodo 2 ha esaurito le connessioni, pertanto il nuovo nodo 17 sceglie il nodo 4 come collegamento diretto.

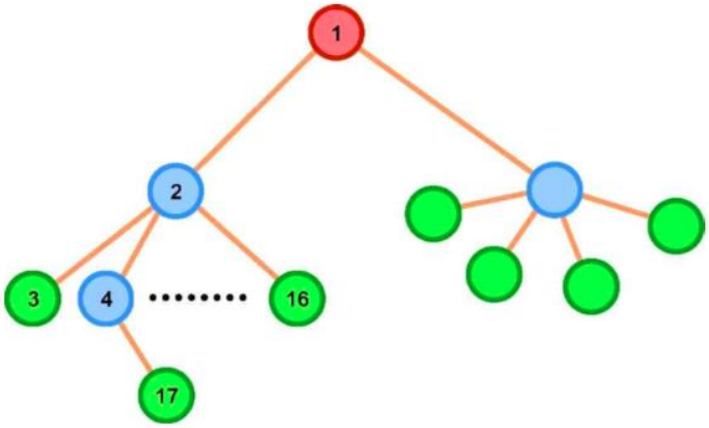


Figura 15 Costruzione della topologia ad albero (Fonte UVAX [48])

All'aumentare dei livelli dell'albero si possono definire anche dei cluster, come mostrato in Figura 16.

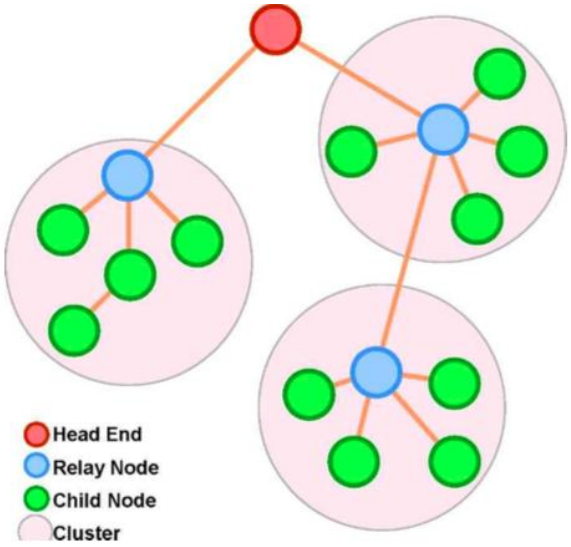


Figura 16 Cluster e Relay Node (Fonte UVAX [48])

I cluster sono gruppi di nodi che si trovano fisicamente vicini. La distanza tra nodi viene misurata in base all'attenuazione della potenza del segnale: questa infatti diminuisce con la distanza. I nodi, in base alla potenza del segnale possono stimare la loro distanza e organizzarsi in cluster.

Ogni nodo ha un *cluster controller* o *relay node*, in grado di comunicare con gli altri cluster e con il concentratore (direttamente o mediante link multi-hop). Il concentratore non appartiene mai ad alcun cluster.

La rete è altamente flessibile, in quanto la configurazione può cambiare al verificarsi di un evento (spegnimento di un nodo, attivazione di un nodo a maggiore capacità, etc). Per mantenere questa flessibilità la topologia della rete è mantenuta attraverso un algoritmo in due fasi.

Quando un nodo si attiva per la prima volta in una rete oppure ha perso la connessione con il *parent*, esegue la prima fase dell'algoritmo di costruzione della rete. In questa fase l'obiettivo del nodo è ottenere il più velocemente possibile la connessione alla rete: in questa fase il nodo vuole trovare un *parent*, senza che questo sia quello ottimo. La scelta, infatti, potrà essere successivamente cambiata e raffinata.

In questa prima fase il nodo riceve dalla rete i pacchetti ACCESS REQUEST PPDU, che vengono inviati dai vicini ogni 30s. In questi pacchetti sono contenuti alcune informazioni:

- indirizzo MAC,
- numero di hop per raggiungere il concentratore,
- numero di porte disponibili,
- guadagno in ricezione.

In base a queste informazioni il nuovo nodo sceglie un *parent* e accede alla seconda fase.

La seconda fase è eseguita periodicamente da tutti i nodi collegati nella rete, in modo da mantenere una topologia ottima. Il nodo che esegue la seconda fase colleziona dalla rete i pacchetti ACCESS REQUEST PPDU e CHANNEL ESTIMATION PPDU, attraverso i quali è possibile conoscere alcuni parametri dei candidati ad essere *parent* ottimi. In particolare vengono forniti i seguenti parametri:

- indirizzo MAC,
- numero di hop per raggiungere il concentratore,
- potenza del segnale,
- rapporto segnale/rumore in ricezione,
- numero di porte disponibili,
- qualità del percorso tra il nodo candidato ad essere *parent* e il concentratore: questa indicazione fornisce una misura dei possibili colli di bottiglia.

Dopo aver collezionato questi dati, il nodo sceglie il suo *parent ottimo*, diventandone *child* e si inserisce nel cluster dei nodi che si trovano fisicamente più vicini.

2.2.9 Sicurezza del sistema Smart Street

La sicurezza di un sistema su rete dipende dalla progettazione della rete e dalla tipologia di servizi che in essa vengono introdotti. In questo paragrafo verranno analizzate alcune scelte progettuali effettuate nell'implementazione della Smart Street oggetto di questo studio per mettere in evidenza i punti di forza e le criticità del sistema.

Uno dei vantaggi più grandi del sistema adottato sta nel fatto che la rete è cablata. Le comunicazioni, infatti, utilizzano come mezzo la linea elettrica. Questa tecnologia, sebbene più soggetta ad attacchi fisici, risulta intrinsecamente più resistente ad attacchi cyber: confrontata con una rete wireless, la rete cablata, organizzata secondo quanto esposto nel precedente paragrafo, risulta più complessa da penetrare. Sarà pertanto più difficile individuare informazioni a partire dai pacchetti presenti sulla rete.

Un ulteriore elemento di sicurezza è rappresentato dalla criptazione. Tutti i dati trasmessi sulla rete sono, infatti, criptati secondo il algoritmo di cifratura a blocchi DES/3DES (Data Encryption Algorithm/Triple Data Encryption Algorithm) per realizzare il totale isolamento tra reti e garantire la privacy degli utenti. Ogni blocco di dati è criptato secondo una chiave generata in maniera casuale secondo l'algoritmo DES. La chiave DES (56 bits) viene a sua volta criptata e trasmessa in testa al blocco. La chiave viene codificata con l'algoritmo 3DES (168 bits), le cui chiavi condivise vengono generate a partire da una stringa configurabile e da un algoritmo di hash.
Questa tecnica di criptazione risulta molto sicura: l'algoritmo 3DES non è attaccabile a forza bruta in quanto utilizza 3 livelli annidati di criptazione (ossia l'algoritmo DES viene applicato 3 volte utilizzando 3 chiavi diverse). Tale algoritmo è considerato tra i più affidabili.
Per quanto riguarda l'algoritmo DES questo è attaccabile per forza bruta, ma la chiave viene sostituita ad ogni blocco, quindi non viene messo a repentaglio tutta la comunicazione ma solo il singolo blocco. I tempi

di latenza introdotti dalla criptazione sono minimi in quanto essa viene realizzata tramite acceleratori hardware, invece che essere implementata via software.
È possibile ridurre i livelli di sicurezza, in quanto la rete può essere configurata senza criptazione dei dati o con livelli diversi di protezione: il CMS consente di abilitare o disabilitare la criptazione a seconda del livello di sicurezza che si vuole mantenere nella rete.

La topologia ad albero della rete sembrerebbe mettere in evidenza una criticità nella robustezza: qualora un nodo *parent* venisse attaccato, tutto il suo sottoalbero perderebbe connettività. Come abbiamo notato, tuttavia, la procedura in due step per la definizione della rete viene richiamata ogni volta che il nodo perde la connettività. La rete, quindi, risulta adattativa e non c'è alcun bisogno di intervento da parte di utenti o gestori della rete stessa.
Il meccanismo con cui le procedure di definizione (o ridefinizione) della rete sono richiamate si basa sul *Node Keep-Alive Mechanism*, per cui ogni nodo controlla continuamente la sua connessione e ripete la procedura di inizializzazione qualora la connessione con il *parent* viene persa. Il nodo concentratore, essendo connesso direttamente all'infrastruttura di comunicazione, deve mantenere la connettività con il default gateway ed, eventualmente, con il DHCP. Qualora la connessione con il DHCP venga persa, esistono procedure di recovery. Per quanto riguarda la perdita di connessione tra il gateway e il concentratore, questa può essere ripristinata generando un allarme remoto.

La rete PLC implementata nello Smart Street utilizza il sistema RADIUS (Remote Authentication Dial In User Service) [50] che fornisce servizi di autenticazione, autorizzazione e accounting centralizzato per computer che si connettono ed utilizzano connessioni di rete.

Il protocollo RADIUS, pur avendo qualche criticità a livello di sicurezza, rappresenta lo standard de facto per l'autenticazione remota. Nel sistema in oggetto di analisi viene utilizzato dai nodi concentratori per controllare i nodi che si possono connettere in maniera sicura sulla rete. Questo protocollo permette di gestire l'accesso alla rete dei modem PLC basandosi sull'indirizzo MAC del dispositivo usando uno o più database.
La connessione tra concentratori e server RADIUS è sicura in quanto utilizza una chiave configurabile e il protocollo PAP.

Il sistema CMS, sebbene utilizzi collegamenti sicuri, rappresenta una vulnerabilità del sistema in quanto soggetto alle vulnerabilità tipo phishing precedentemente illustrate. Queste vulnerabilità non è dovuta alla progettazione del sistema, ma all'utilizzo che di questo se ne fa.

Ulteriori criticità a livello di sicurezza possono essere introdotte nel sistema qualora si definiscano servizi particolari. Una errata progettazione di questi potrebbe rendere il sistema vulnerabile ad attacchi esterni.

2.3 I protocolli di comunicazione degli smart street

2.3.1 Introduzione al protocollo TALQ

Il paradigma definito dalle specifiche TALQ [1] permette la creazione di una infrastruttura basata su enti logici in grado di coordinare un sistema di illuminazione esterno (OLN – Outside Lighting Network) composto da attuatori e sensori identificati come dispositivi logici, tramite un sistema centralizzato (CMS – Central Management System), quale quello mostrato in Fig. 17.

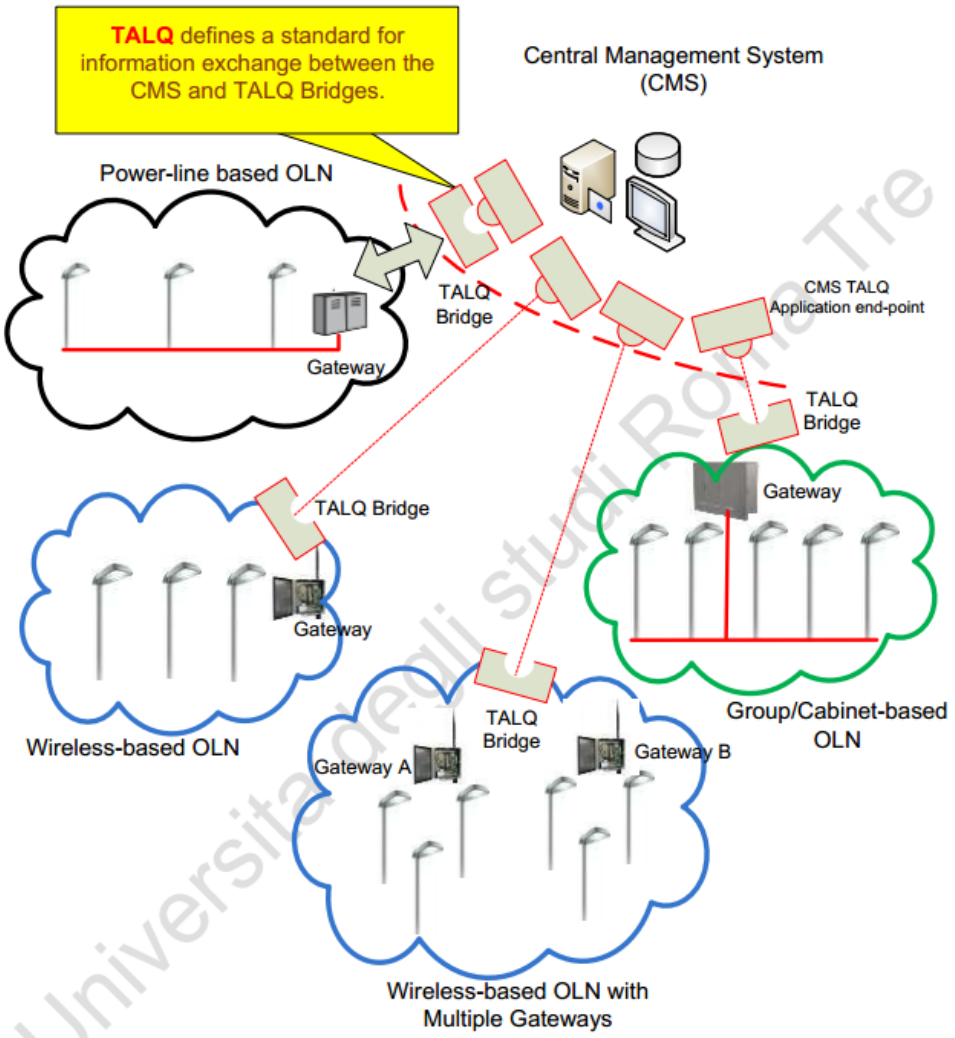


Figura 17 – Paradigma di controllo di un sistema basato su architettura TALQ.

Le specifiche forniscono tutti gli strumenti necessari per la configurazione, il controllo e la gestione degli eventi relativi a un sistema di illuminazione remoto di larghe dimensioni. Il protocollo TALQ, inquadrato all'interno di un modello ISO/OSI standard, ricopre il layer più elevato (Application). Per questo motivo, si appoggia direttamente sugli strati inferiori che garantiscono una corretta comunicazione tra gli elementi interessati all'interno del sistema.
L'implementazione degli elementi inferiori della pila è lasciata allo sviluppatore, in particolar modo per ciò che riguarda lo strato inferiore (PHY) che è fortemente differente a seconda della tipologia di sistema di illuminazione (wired/wireless) presente. Lo studio di queste specifiche è stato necessario nell'ottica

dell'implementazione di un sistema centralizzato per il controllo intelligente di una rete di illuminazione stradale. Il sistema sfrutta una codifica XML e un trasporto tramite protocollo HTTP (HyperText Transfer Protocol) per supportare la messaggistica necessaria al controllo e al *sensing* remoto. La finalità di un sistema di controllo e monitoraggio implementato secondo le specifiche TALQ è di ottenere un sistema rispondente alle seguenti caratteristiche:

- Supporto completo di un sistema messaggistica, tra gli enti interessati, basato su uno standard comune di comunicazione, archiviazione e notifica.
- Layer di messaggistica basato su protocollo HTTP (con possibile espansione a CoAP, nelle versione future dello standard)
- Archiviazione dati in codifica XML (con possibile espansione a EXI e JSON nelle versione future dello standard)

Lo studio del protocollo si è articolato in una analisi della architettura implementata dallo stesso. L'intero sistema è basato su un paradigma di tipo Client-Server. Il ruolo del server è ricoperto dall'entità di controllo centrale, denominata nel protocollo come Central Management System.
Questa unità ha lo scopo di fornire uno strato di Application Protocol con il quale possa interfacciarsi un utente umano, la gestione ad alto livello degli eventi e lo scheduling delle attività. Il protocollo TALQ permette un accesso avanzato ai dati raccolti dai sensori remoti, e per questo motivo, all'interno del CMS, è possibile includere ulteriori algoritmi che gestiscano i dati raccolti tramite protocollo TALQ per ulteriori analisi e azioni di controllo.
Il ruolo del client è ricoperto da una entità di connessione verso i dispositivi finali denominata TALQ Bridge. Questa unità ha lo scopo di gestire le comunicazioni a basso livello tra il CMS e i dispositivi terminali, che possono essere sia sensori sia attuatori. Sia il CMS che il TALQ Bridge supportano l'application layer e il messaging (HTTP). Non essendo lo strato fisico definito dallo standard, sono stati considerati all'interno di questo studio diversi strati fisici e protocolli di comunicazione adatti alle diverse esigenze che possono nascere nella comunicazione in ambienti complessi come quello dell'illuminazione stradale. La criticità delle informazioni e dei segnali di controllo usati per il sistema di illuminazione hanno richiesto l'implementazione di un paradigma di comunicazione che supportasse una codifica sicura. A questo scopo, sono state approfondite le implementazioni della tecnologia HTTP over TLS (Transport Layer Security) supportate dalle specifiche TALQ.
L'ultimo elemento della catena implementata dalla architettura TALQ sono i dispositivi logici (Logic Device) Ogni dispositivo logico può essere associata ad uno o più dispositivi fisici di tipo sensore o attuatore. Al fine di caratterizzare le potenziali applicazioni tecnologiche supportate dalle specifiche TALQ in termini di dispositivi fisici compatibili, sono stati studiati gli elementi più importanti considerati nella messaggistica tra i dispositivi logici ed il TALQ Bridge. È stato verificato che le specifiche permettono un monitoraggio di grandezze fisiche di estremo interesse per una gestione *smart* del processo di illuminazione. Parametri come assorbimento di potenza, tensione, corrente, fattore di potenza, e tempo di operatività sono implementati nelle strutture XML di immagazzinamento dati.
Le stesse grandezze sono riconosciute all'interno degli algoritmi di riconoscimento e gestione degli eventi implementati ad alto livello all'interno del CMS.

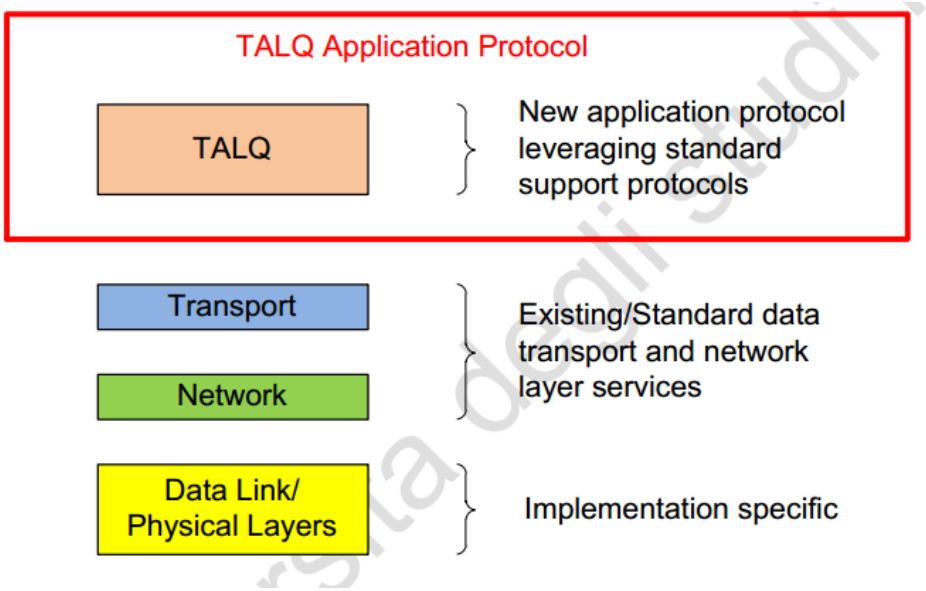


Figura 18 – Inquadramento del protocollo TALQ all'interno di uno stack ISO/OSI

Il protocollo prevede, per le suddette grandezze, un sistema avanzato di data-logging. Questa caratteristica è critica ai fini dello sviluppo del progetto in esame, essendo l'analisi delle serie temporali di dati su base statistica un elemento fondamentale per la stima e il controllo intelligente delle risorse energetiche. Per verificare l'effettiva compatibilità del protocollo TALQ con le specifiche realizzative richieste dal progetto in esame, è stato effettuato uno studio delle principali funzionalità implementate all'interno del sistema di gestione CMS. Il sistema supporta una messaggistica sincrona e asincrona bidirezionale tramite la quale è possibile una interazione con dispositivi anche in scenari estremamente ramificati. Ciò è possibile grazie alla presenza del TALQ Bridge di interfaccia verso le unità terminali. Il CMS e il TALQ Bridge supportano un servizio di scoperta e configurazione guidata dei dispositivi logici (configuration service).
Questa caratteristica è utile al fine di gestire un sistema dinamico in cui la architettura di sensori e attuatori può dover essere modificata sia per espansione del sistema sia per motivi di manutenzione. Particolare attenzione è stata posta in questo studio nei riguardi del servizio relativo al controllo degli attuatori preposti all'illuminazione (Lighting Control).
Questo servizio permette di impostare il livello di illuminazione dei vari attuatori secondo diverse metodologie. È possibile, dal CMS, impostare in maniera diretta il livello di illuminazione (Override) o sfruttare dei sistemi complessi basati su programmi (Scheduling). Questa seconda funzionalità permette di assegnare ai dispositivi logici relativi agli attuatori di illuminazione delle tabelle di funzionamento che possono essere basate su calendari e programmi di controllo.
Il controllo su base temporale (Calendar) permette di implementare un sistema di illuminazione che possa modulare il proprio funzionamento su base periodica. Questo sistema è utile al fine di realizzare un controllo intelligente che moduli il consumo di potenza risparmiando energia in periodi a bassa richiesta. Il controllo su base programma (Control Program) permette di aggiungere un set di regole dinamiche che integrano il calendario. Queste regole sfruttano segnali che provengono sia dal CMS (tramite, per esempio, un controllo su database relativi alle condizioni metereologiche) sia dalla sensoristica identificata dai relativi dispositivi logici. In questo paradigma, le specifiche TALQ definiscono inoltre un insieme di eventi, intesi come la concomitanza di differenti tipologie di segnali provenienti dai dispositivi logici. L'implementazione di una gerarchia di eventi per la gestione di malfunzionamenti, comportamenti anomali ed emergenze funzionali è di primaria importanza per un sistema critico come quello dell'illuminazione stradale. Complessivamente, dallo studio fatto è emerso che la flessibilità offerta dalle specifiche TALQ permette il completo interfacciamento tra algoritmi di gestione intelligente del consumo energetico e sistemi di illuminazione realizzati ad-hoc.

2.3.2 Messaggi - Descrizione

Il protocollo TALQ definisce, a livello di Application Data (XML) tre tipologie di messaggi:

- *request message*: proveniente dal TALQ Bridge e diretto al CMS.
- *response message*: proveniente dal CMS e diretto al TALQ Bridge.
- *notification message*: messaggistica asincrona dal CMS al TALQ Bridge.

I contenuti di questi messaggi è definito da tipi di dato complessi in XML che descrivono attributi ed elementi di dispositivi e servizi. I messaggi inviati sono incapsulati in *container documents* che permettono l'invio di messaggi multipli, della stessa tipologia, in un unico messaggio. Sono definite due tipologie di *container documents*:

- <requestDoc>: Questo *container document* raggruppa in se uno o più messaggi di *request* inviati dal TALQ Bridge al CMS. Ogni elemento è contraddistinto da una tag individuale <request>.
- <responseDoc>: Questo *container document* raggruppa in se uno o più messaggi di *response* inviati dal CMS al TALQ Bridge come risposta all'invio di un <requestDoc>. Ogni elemento è contraddistinto da una tag individuale <response>.

Entrambi i *container documents* sono definiti come elementi a livello *top*. I singoli messaggi raggruppati nel *container document* sono definiti dal protocollo, e per ciascuno di essi, un attributo "*xsi:type*" ne descrive la tipologia. Per permettere al TALQ Bridge di associare i contenuti del *responseDoc* alle query effettuate con il *requestDoc*, sono contemplati due metodologie di ordinamento: le risposte possono essere fornite con lo stesso ordine delle query, o per ciascun elemento del *responseDoc* può essere inserito un attributo "*reqOrd*" che ne descriva l'ordine. Allo scopo di identificare univocamente il TALQ Bridge di riferimento per l'invio e la ricezione dei *requestDoc* e *responseDoc*, in entrambi i documenti è incluso, nel *top level*, un attributo chiamato "*bridgeAddress*", che corrisponde al TALQ Address del dispositivo logico che implementa la funzione di TALQ Bridge. Questo indirizzo è creato ed assegnato dal CMS ai vari TALQ Bridge che sono presenti nel sistema di illuminazione durante la procedura di *bootstrap*. In assegnza di un indirizzo correttamente assegnato, il TALQ Bridge invierà al CMS una richiesta di notifiche *getNotifications* al fine di ottenere l'assegnazione del *bridgeAddress.* La risposta del CMS in questo caso sarà un messaggio di tipo *synchronizeBridge* tramite il quale viene assegnato, al bridge, il proprio indirizzo. Nel caso venga inviato, da una entità qualsiasi, un messaggio privo dell'attributo "bridgeAddress", la risposta sarà sempre un codice di errore HTTP 400. Il *bridgeAddress* è un caso particolare di indirizzo assegnato, nel particolare caso, ai TALQ Bridge. Tutte le entità e le risorse in un paradigma TALQ sono identificate da un indirizzo TALQ. Gli indirizzi TALQ sono costruiti sul seguente formato:

<address-domain>:<entity-address-part>

Il campo <address-domain> definisce la tipologia di indirizzo TALQ, e permete di distinguere dispositivi logici (dev), gruppi (GRP) o risorse indirizzabili in una ONL. Esempi di risorse indicizzabili sono i calendari, i programmi di controllo e i *data logger*. Una tabella riassuntiva delle tipologie di indirizzi contemplati nel protocollo è riportata in Tabella 1. Come è possibile vedere, alcuni indirizzi sono assegnati (*owner*) dal TALQ Bridge, altri dal CMS. Il bridgeAddress è un caso particolare (ed unico) di indirizzo assegnato a Logical Device per il quale l'owner è il CMS. Per permettere una maggiore leggibilità, è seguita la convenzione di usare lettere maiuscole per i domini delle entità assegnate dal CMS, e lettere minuscole per le entità assegnate dai TALQ Bridge.

Tabella 1 – Tipologie di indirizzi TALQ disponibili nel protocollo: dominio, entità responsabile dell'assegnazione, descrizione.

Domain	Owner (Assigns the Address)	Description
dev	TALQ Bridge	Logical device[1]
GRP	CMS	Group
SCN	CMS	Scene
cls	TALQ Bridge	Device class
LMP/lmp	CMS/TALQ Bridge	Lamp type
CPR	CMS	Control program
CAL	CMS	Calendar
DLG	CMS	Data logger

[1] Il dominio *dev* presenta una eccezione nel caso del TALQ Bridge, per il quale l'*owner* è il CMS.

Il sistema di messagistica implementato prevede una robustezza nei confronti di errori di trasmissione. Ogni messaggio di notifica prevede un numero di identificazione sequenza, definito come attribuo xml seq. Il numero sarà incrementato ad ogni successivo messaggio fino ad overflow (2^{23} -1).
Nel caso un TALQ Bridge ricevesse un messaggio con numero di sequenza sbagliata lo ignorerà ed invierà al CMS un request con l'ultimo numero di sequenza ricevuto. Il CMS, di risposta, riprenderà ad inviare messaggi con il numero di sequenza direttamente successivo rispetto a quello inviato dal TALQ Bridge.

2.3.3 Messaggi - Supporto

Tutti i messaggi del protocollo TALQ sono mappati su requests HTTP con protocollo di trasporto TCP. Questa scelta garantisce l'uso di un protocollo di trasporto efficace, robusto, ed ampiamento supportato da numerose piattaforme.
La comunicazione deve essere implementata su request HTTP-POST. La configurazione è diversa dal classico HTTP usato nel web browsing, in quanto la maggior parte del traffico in un sistema TALQ avviene da client (identificato dalle ONLs) a server (CMS). Il server, dopo aver ricevuto la *request*, esegue la azione richiesta e risponde immediatamente inerentemente allo stato della azione. In Fig. 19 è possibile vedere una normale comunicazione *request-response* basata sul metodo HTTP-POST.
Allo scopo di permettere la comunicazione in caso di eventi particolari, il sistema prevede una messagistica asincrona da parte del CMS basata su *HTTP long polling*.
Le *request* HTTP possono essere terminate per numerose ragioni diverse da una *response* del CMS. In questo caso il TALQ Bridge invierà al CMS un *long poll request* (*getNotifications*) dopo un timeout configurabile tramite l'attributo "*pollTimeout*" definibile sul TALQ Bridge. La scelta di questo attributo deve essere fatta sulla base di un compromesso tra la latenza nelle comunicazioni Server-Client e le risorse necessarie per la trasmissione dei *long polling*.

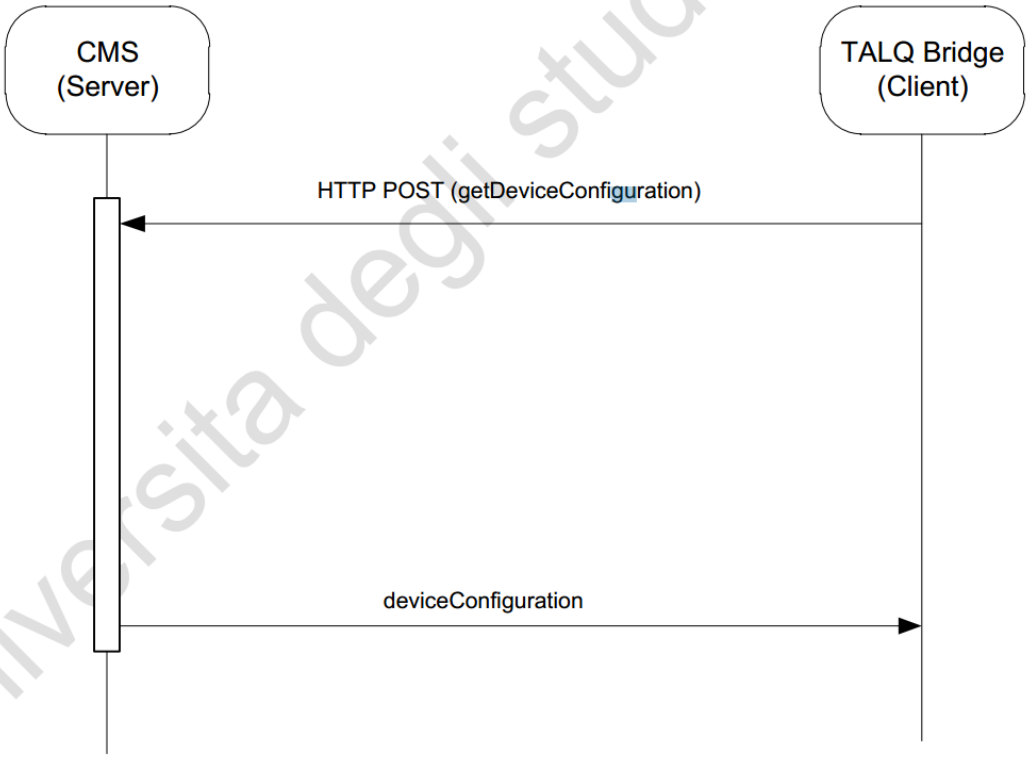


Figura 19 – Rappresentazione schematica di una tipica comunicazione HTTP-POST di tipo *request-response* tra un CMS (Server) e un TALQ Bridge (Client)

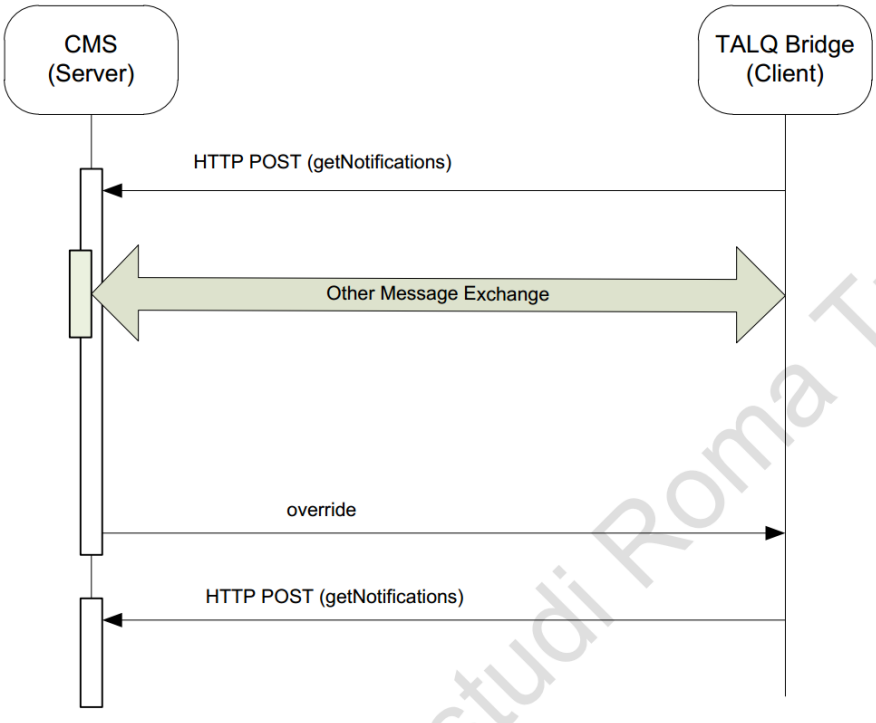


Figura 20 – Rappresentazione schematica di una tipica comunicazione HTTP long polling.

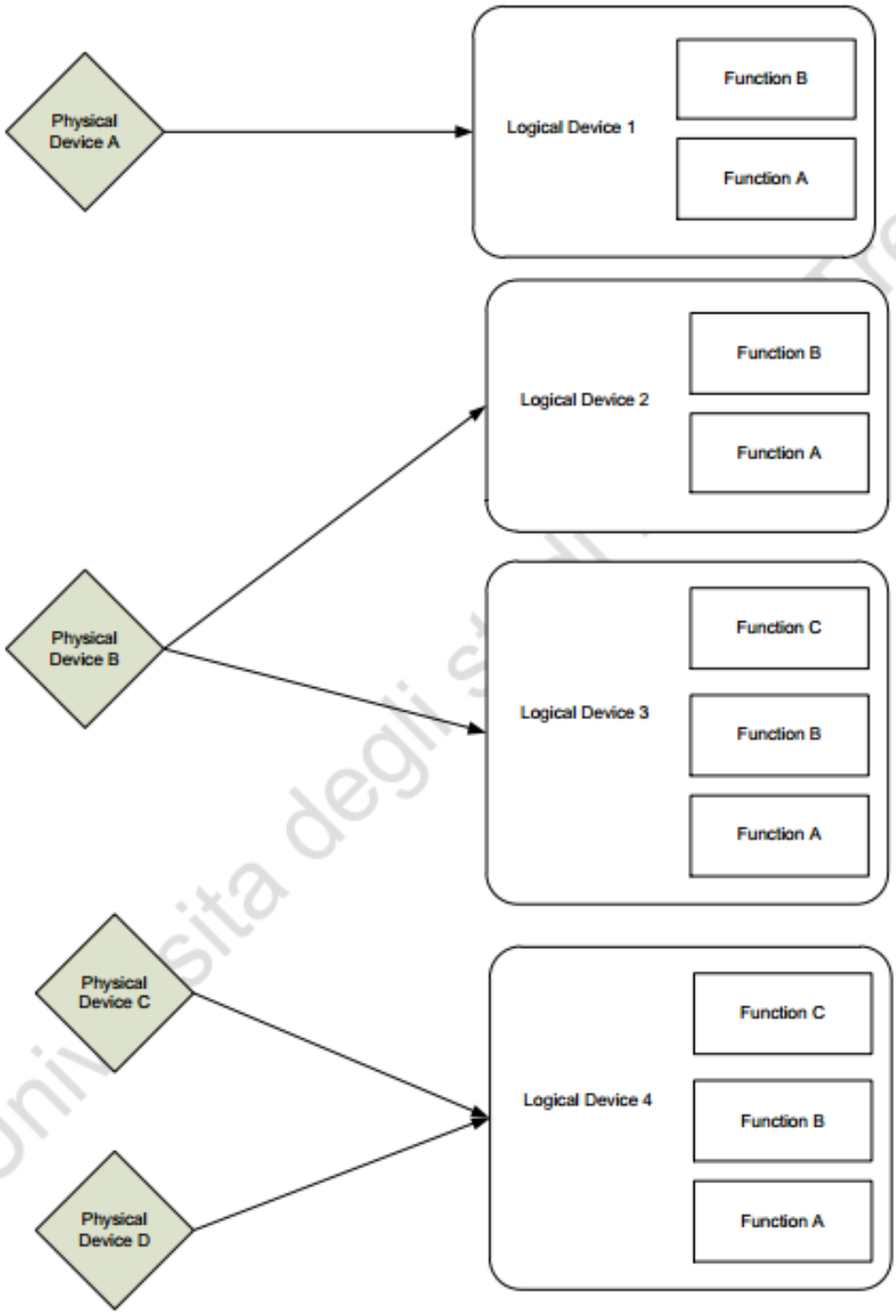


Figura 21 – Gerarchia dei dispositivi fisici, logici e delle funzioni.

2.3.4 Modello Dati

Le specifiche TALQ adottano dei modelli di dato per i quali le risorse di sistema e le funzionalità ad esse associate sono raggruppate in *functions* entro dispositivi logici (e.g. attuatori di illuminazione, sistemi di misura e sensoristica). Questo modello è usato per l'implementazione delle applicazioni terminali di TALQ. Una *function* TALQ consiste in un insieme di attributi ed eventi, che descrivono e supportano una specifica funzionalità. Un dispositivo logico è la rappresentazione logica di un dispositivo, e questo implementa a sua volta numerose funzioni. Un esempio di dispositivo logico può essere un controllore di illuminazione, che al

suo interno integra funzioni di impostazione di luminosità, monitoraggio e controllo di grandezze misurabili. Un dispositivo logico associato ad un armadio di controllo può contenere funzioni relative all'attivazione e spegnimento delle luci e al monitoraggio dei consumi. In termini gerarchici, un singolo dispositivo logico può integrare al suo interno più funzioni, più dispositivi logici possono essere associati ad un singolo dispositivo fisico, e più dispositivi fisici possono essere associati allo stesso dispositivo logico. La piattaforma di sviluppo è flessibile, e si presta ad essere adattata alle diverse tipologie di prodotti presenti sul mercato.

2.3.5 Funzioni

Una funzione TALQ (*TALQ function*) consiste in un insieme di attributi ed eventi che descrivono una funzionalità legata ad un dispositivo logico. Modificando gli attributi della funzione, inerentemente alla tipologia di servizio TALQ, è possibile configurare, pilotare e gestire il dispositivo logico sul quale la funzione è implementata. Nelle specifiche TALQ sono definite, correntemente, le seguenti funzioni:

- Basic
- TALQ Bridge
- Communication
- Lamp Actuator
- Lamp Monitor
- Electrical Meter
- Photocell
- Light Sensor
- Binary Sensor
- Generic Sensor
- Time

Ciascuna funzione può essere dotata di differenti attributi, che ricadono nelle seguenti categorie:

- Configurazione (*configuration*): attributi che descrivono le capacità e le caratteristiche del dispositivo. Raggruppano tutti gli attributi di configurazione che determinano le modalità di funzionamento per i dispositivi. Gli attributi di configurazione forniscono informazioni che possono essere usate per configurare in maniera ottimale il sistema. Gli attributi possono essere aggiornati dal CMS in corso d'opera o durante la fase di inizializzazione del sistema.
- Operativi (*operational*): attributi che possono essere manipolati attivamente da un servizio TALQ per avere un controllo diretto del dispositivo in funzione di una richiesta effettuata dal CMS in corso d'opera. Un esempio significativo è la luminosità desiderata da un attuatore luminoso.
- Misura (*measurements*): attributi che forniscono dati inerenti le prestazioni del dispositivo. Le quantità di interesse in questo caso possono, per esempio, essere energia, potenza, tensione, corrente.
- Stato (*status*): attributi associabili a specifici eventi che possono essere impiegati per indicarne lo stato. Sono considerati opzionali dallo standard, e in generale, non descritti per evitare ridondanza.

Segue una breve descrizione delle principali funzioni di interesse per lo studio effettuato.

BASIC

La funzione Basic definisce le proprietà legate alle risorse fisiche cui sono associati i dispositivi logici. Queste proprietà sono molteplici, tra le quali, l'identificazione (assetId) e la posizione. Un esempio di risorsa fisica è un palo della luce su cui son montate due o più luminarie. Ogni luminaria può essere modellata come un dispositivo logico indipendente, ma tutte associate alla stessa risorsa fisica. La funzione Basic include anche gli attributi che si applicano ai dispositivi logici in maniera integrale e non alle specifiche funzioni implementati dagli stessi. Esempi possono essere il numero seriale, la tipologia di hardware, la versione e la data di installazione. Gli attributi previsti dal protocollo per la funzione Basic sono riportati nella tabella seguente:

Tabella 2 – Attributi funzione Basic

Id Attributo	Tipo	Descrizione
assetId	xs:token	Identificativo della risorsa fisica associata. Se più dispositivi hanno lo stesso assetId significa che appartengono alla stessa risorsa.
latitude	float	Latitudine WGS84
longitude	float	Longitudine WGS84
altitude	float	Metri sopra il livello del mare
serial	xs:token	Numero seriale del dispositivo
hwType	xs:token	Tipo di hardware del dispositivo
swType	xs:token	Tipo di software del dispositivo. Utile nel caso lo stesso hardware supporti diverse versione di firmware.
installationDateTime	dateTime	Data e ora dell'installazione del dispositivo

Gli eventi supportati dalla funzione Basic sono riportati nella seguente tabella:

Tabella 3 – Eventi funzione Basic

EventType	Descrizione
deviceReset	Il dispositivo fisico contenente il dispositivo logico è stato riavviato.
softwareUpdating	Il software del dispositivo è in corso di aggiornamento
hardwareUpdating	L'hardware associato al dispositivo logico è stato aggiornato
batteryMode	Dispositivo operante con alimentazione a batteria.
installationMode	Il dispositivo è in fase di installazione.
maintenanceMode	Il dispositivo è in fase di manutenzione, dove la manutenzione include azioni riguardanti hardware o software.
cabinetDoorOpen	Segnale di apertura armadio.
batteryShutdown	Il dispositivo si è spento a causa di esaurimento batteria
locationUpdated	Variazione della posizione del dispositivo.

TALQ Bridge

La funzione TALQ Bridge include gli attributi necessari a permettere la comunicazione tra il CMS e l TALQ Bridge. Per ogni OLN, è presente una e una sola istanza della funzione TALQ Bridge. Sono supportati i seguenti attributi:

Tabella 4 – Attributi funzione TALQ Bridge

Id Attributo	Tipo	Descrizione
cmsUri	AnyUri	URI di base per la comunicazione TALQ che permetta al TALQ Bridge di accedere al CMS. Deve essere un URI assoluto. Altri URI relativi di accesso al CMS possono usare questo indirizzo come base.
bootstrapComplete	Boolean	Flag impostata dal CMS a TRUE per indicare un processo di bootstrap completo del TALQ Bridge.
pollTimeout	unsignedInt	Tempo, contato dopo aver inviato un *getNotifications* (long poll), dopo il quale la richiesta non è più valida e viene reiterata dal TALQ Bridge.
retryPeriod	float	Tempo oltre il quale il TALQ bridge ritrasmette un messaggio per il quale è attesa una risposta non ricevuta.
vendor	xs:token	Produttore
pkgUrl	AnyUri	URL che punta alla posizione di pacchetti scaricabili. Usato dal servizio di trasferimento dati.
currentReleaseId	xs:token	Versione corrente.
lastNotificationSeq	Talq:Sequence	Numero sequenziale dell'ultimo messaggio di notifica ricevuto dal TALQ Bridge

Non sono definiti eventi specifici per la funzione TALQ Bridge. Possono essere tuttavia generati, da esso, eventi specifici relativi a servizi.

Lamp Actuator

La funzione Lamp Actuator include gli attributi relativi al cotrollo dell'illuminazione e rappresenta la più piccola unità con scopo di controllo. La funzione può tuttavia controllare più di un dispositivo fisico, e può essere legata al pilotaggio di combinazioni di lampade e apparati di controllo in maniera centralizzata. Può essere presente al massimo un Lamp Actuator per dispositivo logico, che può quindi corrispondere a più carichi. Se l'applicazione richiede di controllare indipendentemente più attuatori luminosi, essi saranno modellati come singoli dispositivi logici. Gli attributi supportati sono riportati nella tabella seguente. Principalmente sono attributi di tipo operativo. Segue una tabella degli attributi più importanti della funzione:

Tabella 5 – Attributi funzione Lamp Actuator

Id Attributo	Tipo	Descrizione
lampTypeId	TALQ Address	Indirizzo TALQ di un lampType esistente.
standbyMode	Enumeration	Definisce il comportamento dell'attuatore quando il livello di output è zero. OFF: il livello luminoso è zero e non viene inviata alimentazione ai sistemi di controllo; ON: il livello luminoso è zero ma i sistemi di controllo sono alimentati.
cloEnabled	Boolean	Determina se è previsto per il dispositivo una correzione di tipo Constant Light Output (CLO). Il fattore di correzione è usato per compensare l'invecchiamento della lampada. Il fattore è specificato all'interno del lampType.
maintenanceFactorEnabled	Boolean	Abilitazione della compensazione da manutenzione.
maintenanceFactor	integer (0 to 100%)	Fattore aggiuntivo di correzione (addizionabile al CLO) che corregge la luminosità in caso di manutenzione. Il fattore 0 corrisponde alla lampada appena pulita.
mainteinancePeriod	unsignedInt	Periodo oltre il quale il fattore di mainteinance raggiunge il 100% (e.g. lampada completamente sporca). Si assume una progressione lineare.

defaultLightState	LightState	Imposta l'output di default per l'attuatore luminoso. Questo stato è usato se nessun altro comado è attivo, e di default, è impostato al 100%.
calendarID	TALQ Address	Indirizzo TALQ del calendario che controlla l'attuatore luminoso (*cal:CalendarID*). In assenza di questo valore, o nel caso referenziasse un indirizzo non valido, il comportamento è impostato dalla OLN.
targetLightCommand	LightCommand	Ultimo comando impostato sull'attuatore luminoso.
feedbackLightCommand	LightCommand	Comando attuale per l'attuatore luminoso.
actualLightState	Light State	Questo comando rappresenta lo stato fisico dell'attuatore luminoso nella maniera più completa, includendo fattori di correzione come il CLO e il maintenance. Può essere calcolato o misurato, a seconda dell'implementazione del dispositivo fisico.

Sono supportati diversi eventi, tra i quali:

Tabella 6 – Eventi funzione Lamp Actuator

EventType	Descrizione
lightStateChange	Lo stato dell'attuatore luminoso è cambiato. Questo evento è segnalato al CMS all'interno del data log.
invalidCalendar	L'attuatore luminoso è stato configurato con un calendario non implementabile
invalidProgram	L'attuatore luminoso è stato configurato con un programma non implementabile

2.3.6 Dispositivi Logici

La flessibilità del paradigma TALQ porta a definire i dispositivi logici come l'insieme di funzioni da essi implementati. L'associazione del dispositivo logico al dispositivo fisico in questo modo è estremamente semplice e flessibile. Tipici dispositivi logici che possono essere implementati sono: controller di luminosità, armadio di controllo, multimetro, TALQ Bridge, interfaccia di comunicazione. I dispositivi TALQ sono organizzati in classi (o tipi) cui vengono assegnati dal TALQ Bridge in ottemperanza alle specifiche del protocollo. La definizione della classe di un dispositivo è usata dal servizio di configurazione per effettuare un annuncio delle configurazioni dei dispositivi con il minimo over head di trasmissione possibile. Una volta che una classe di un dispositivo è stata annunciata, altre istanze possono essere annunciate senza ripetere completamente la trasmissione della descrizione del dispositivo. Ogni dispositivo logico è identificato dal suo indirizzo univoco, con dominio *dev* (riservato ai dispositivi logici). Per accedere agli attributi interni al dispositivo è necessario appenderne l'estensione (funzione e attributo) all'indirizzo univoco, per esempio:

dev:<dev-address>[:<function id>[:<attribute id>]]

L'identificatore *function id* definisce una particolare istanza di una funzione all'interno di un dispositivo logico. La *function id* è assegnata dal TALQ Bridge, che gestisce i dispositivi logici. Va notato che gli indirizzi TALQ possono essere usati per riferirsi a singoli target. Se per esempio è necessario riferirsi a più funzioni, sarà necessario usare una lista di indirizzi.
Un esempio di dispositivo logico è mostrato in Fig. 22.

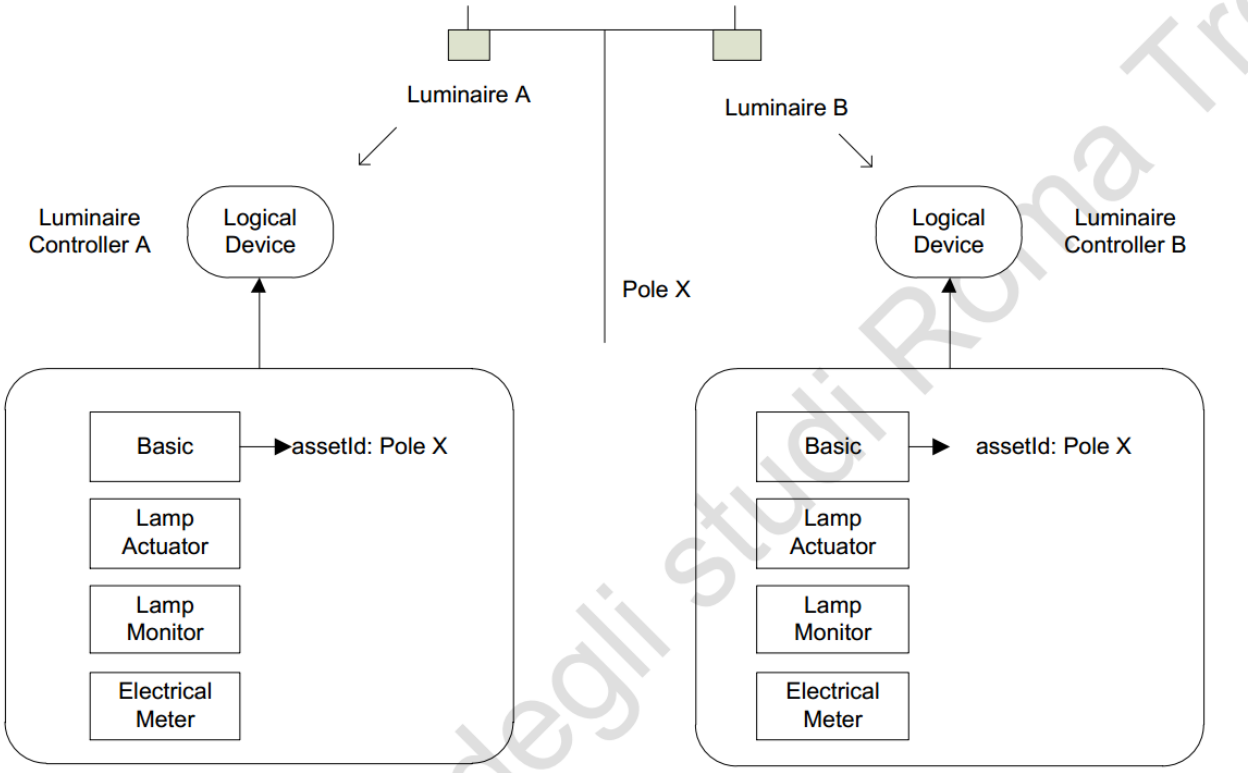


Figura 22 – Esempio pratico di due dispositivi logici assegnati allo stesso asset fisico.

Due lampade (A e B) connesse allo stesso palo sono modellate come due dispositivi logici indipendenti (denominati Luminare Controllers). Ciascuno di essi può essere pilotato indipendentemente ed è dotato delle funzioni Basic, Lamp Actuator, Lamp Monitor e Meter.

2.3.7 Servizi

I servizi TALQ definiscono i messaggi scambiati e il comportamento delle applicazioni necessario per manipolare gli attributi delle funzioni. Sono definiti i seguenti servizi:

- Notifica (*notification*): servizio che definisce i messaggi asincroni dal CMS al TALQ Bridge sfruttando il meccanismo di comunicazione *HTTP long polling.*
- Configurazione (*configuration*): servizio che definisce i processi di configurazione e amministrazione del sistema. Sono inclusi in questo servizio i processi di bootstrap, di discovery e di update.
- Controllo illuminazione (*lighting control*): particolare servizio di controllo dedicato esclusivamente ai dispositivi di illuminazione. Gestisce calendari e impostazioni dirette (*override*)
- Dati On-Demand (*on demand data request*): servizio che gestisce il processo request-response per accedere ad informazioni operative, di misura e di stato dei dispositivi appartenenti ad una OLN on-demand.
- Gestione Gruppi (*group management*): servizio di gestione e amministrazione di gruppi di dispositivi logici e gruppi di funzioni.
- Trasferimento Dati (*data package transfer):* servizio di gestione trasmissione verso il TALQ Bridge per supportare dati specifici del produttore della OLN.

Il servizio di Lighting Control, in particolare, è quello più vicino all'apparato fisico che implementa il sistema di illuminazione. Il sistema di controllo Override, in particolare, fornisce al CMS un controllo diretto degli attuatori luminosi, permettendo un controllo centralizzato in grado di aggirare le impostazioni di calendari e programmi. In assenza di controllo diretto, esiste un controllo schedulato che permette una impostazione dell'output luminoso in funzione del tempo e delle letture dei sensori. Il servizio deve essere supportato sia dal CMS sia dal TALQ Bridge. Indubbiamente, possono esistere delle OLN prive di attuatori luminosi controllabili (e.g. dotate esclusivamente di sensori). In questo caso, il servizio è comunque di fondamentale importanza lato TALQ Bridge per garantire la possibilità di monitorare eventi come la sostituzione di un modello di lampada, o un processo di manutenzione. Nell'ambito del servizio è definita l'entità Lamp Type, all'interno della quale sono elencati una serie di parametri che permettono di caratterizzare completamente un attuatore luminoso. Un riferimento a una tipologia di lampada è incluso nelle funzioni Lamp Actuator e Lamp Type, essendo la tipologia di lampada un parametro necessario al corretto funzionamento di queste funzioni. Attributi interessanti specificati nell'entità Lamp Type includono:

- Potenza della lampada.
- Controllo mediante PWM, switch, DALI (Digital Addressable Lighting Interface) o altri metodi.
- Massima e minima tensione di controllo.
- Soglie operative di corrente minima e massima.
- Massimo periodo operativo (ore)
- Pendenza curva lumen/potenza
- Profilo di correzione degrado luminoso della lampada.

La luminosità istantanea della lampada è impostata, nella funzione Lamp Actuator, impostando un attributo con tipo Light State, che indica un livello percentuale della massima potenza luminosa erogata dalla lampada.

2.3.8 Eventi

Il protocollo TALQ prevede il supporto per gli eventi (*events*). Gli eventi sono riportati dalla OLN al CMS per semplificare il processo di gestione e diagnosi. Gli eventi forniscono informazioni inerenti a situazioni o condizioni di utilità nella gestione della OLN. Ciascun evento viene identificato da un tipo (*eventType*) e da una descrizione *(Description)*. Gli eventi sono comunicati al CMS tramite dei log, gestiti dal servizio di Data Collection. La breve descrizione associata all'evento può essere configurata per supportare diversi linguaggi: è possibile specificare la descrizione dell'evento in più lingue, tuttavia non è possibile inserire più descrizioni per la stessa lingua. Si definiscono, nel protocollo, tre tipologie di eventi:

- Generico (*generic*): questa tipologia di evento è generalibile da tutte le tipologie delle funzioni TALQ.
- Specifico di Funzione (function specific): Evento relativo a una funzione particolare.
- Specifico di Servizio (service specific): Eventi relativi a un particolare servizio. In questo caso gli eventi sono definiti come parte del servizio stesso, e dove previsto, gli eventi saranno annunciati come opzioni specifiche del servizio.

Per ogni evento è quindi specificata una sorgente, che può essere un dispositivo logico o una funzione definita all'interno del dispositivo logico. Gli eventi associati ai servizi saranno sempre generati dal TALQ Bridge, o in sua vece, dal dispositivo logico che implementa la funzione di TALQ Bridge. È anche possibile implementare attributi di stato, che possono essere usati come sorgenti per specifici eventi. Gli eventi generici definiti nel protocollo TALQ sono i seguenti:

- invalidAddress: riferimento ad un indirizzo non valido.
- invalidConfig: Il CMS ha fornito un valore non valio per uno o più attributi di configurazione presenti in un dispositivo logico.
- deviceMalfunction: un dispositivo fisico, associato a uno o più dispositivi logici, ha avuto un malfunzionamento. Questo è un evento generale di segnalazione che deve essere usato in assenza di eventi più specifici e descrittivi.
- functionError: un errore è avvenuto all'interno di una funzione. Analogamente al caso precedente, è un segnale di errore generico da usarsi in assenza di segnalazioni più specifiche.
- null: evento nullo, è possibile usarlo per descrivere il log eventi come vuoto.

2.3.9 Sincronizzazione Dati

La sincronia tra server e client in un sistema complesso come quello descritto dal protocollo TALQ è di estrema importanza. Sia il CMS sia i TALQ Bridge devono avere una visione consistente ed uniforme dei dispositivi logici e delle entità gestite dai servizi, quali calendari, programmi di controllo e tipologie di sistemi di illuminazione. Per semplificare la sincronizzazione tra le varie entità, il protocollo prevede che le entità posano essere create e modificate esclusivamente da una entità gerarchicamente superiore che viene chiamata *owner*. Le due tipologie di entità dotate di questi privilegi sono il CMS ed il TALQ Bridge. Le tipologie di entità, e il loro owner, sono specificate nello standard. Nel caso avvenga una creazione o modifica, l'entità *owner* comunica alle altre entità l'azione effettuata, ma nessuno dei riceventi del messaggio può operare ulteriori modifiche a riguardo. Per ogni tipologia di entità solo un *owner* è definito con l'eccezione del lampType, che può essere modificato sia dal CMS sia dal TALQ Bridge.

2.3.10 Gestione delle entità

Le entità sono gestite all'interno del contesto corrispondente ai servizi loro adibiti. Questa sezione descrive l'insieme di funzioni adottate per gestire le entità tramite i servizi. Per ogni tipo di entità sono definiti dei messaggi e contenuti specifici. Il nome del messaggio riflette la tipologia di entità. I seguenti messaggi sono adibiti a gestire le entità appartenenti (*owned*) dal CMS:

- *<entity> Changed:* Messaggi di notifica inviati dal CMS al TALQ Bridge per informare della modifica a carico di una entità. Questo messaggio contiene l'indirizzo TALQ dell'entità che è stata modificata. Il tipo xml usato per questo messagio è <EntitytChangedNotification>.

- *get<entity>*: Messaggio di richiesta dal TALQ Bridge al CMS per ottenere le modifiche effettuate su una entità. Il TALQ Bridge userà questo messaggio come risposta ad un messaggio *<entity>Changed* se la notifica indica che la configurazione della <entity> non è aggiornata nel TALQ Bridge, o quando una nuova entità non ancora nota viene scoperta. Il messaggio di richiesta viene ripetuto fino ad ottenere una risposta di tipo <entity>. Il messaggio contiene l'indirizzo (o gli indirizzi) TALQ dell'entità richiesta dal TALQ Bridge. Il tipo xml usato per questo messaggio è <EntityRequest>.

- *<entity>:* Messaggio di risposta da parte del CMS al TALQ Bridge, contenete i dati richiesti. Può essere un aggiornamento totale o parziale. Il messaggio contiene, a tutti gli effetti, una entità. I dati effettivi dell'entità sono descritti, per ogni messaggio, all'interno delle sezioni relative ai servizi. Generalmente, una entità contiene l'indirizzo TALQ e l'attributo "seq" che ne indica il numero di sequenza, e l'attributo "modifier" che ne descrive la completezza dei dati inclusi.

Per esempio, il servizio Group Management definisce tre tipi di messaggi per gestire i gruppi: *groupChanged*, *getGroup*, *group*. In figura viene rappresentato, schematicamente, il processo comunicativo che avviene tra CMS e TALQ Bridge per gestire una entità appartenente al CMS.

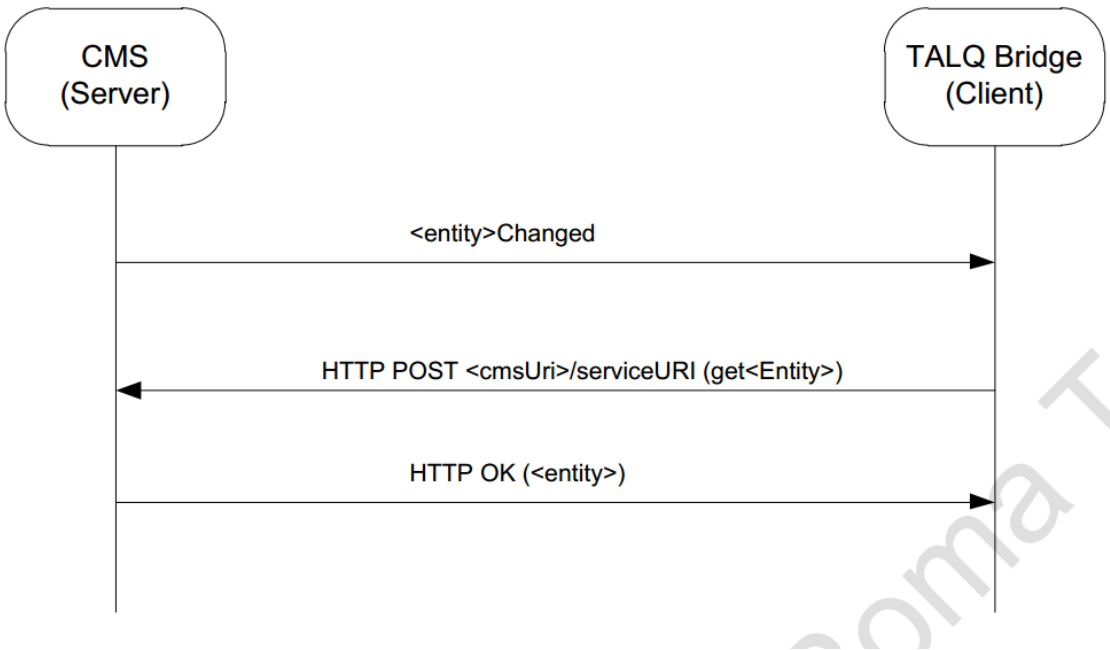


Figura 23 – Comunicazione tra TALQ Bridge e CMS per gestire una entità appartenente al CMS. Il CMS comunica l'avvenuta modifica di una entità, il TALQ Bridge risponde richiedendo la descrizione delle modifiche, il CMS conclude la comunicazione trasmettendo lo stato attuale dell'entità.

Le entità appartenenti al TALQ Bridge sono gestite dai seguenti tipi di messaggio:

- *refresh<Entity>*: Messaggi di notifica dal CMS al TALQ Bridge per richiedere al TALQ Bridge di ritrasmettere i dati relativi ad una specifica entità. Questo messaggio è usato dal CMS per richiedere informazioni nel momento in cui una nuova entità a lui sconosciuta viene scoperta. Il CMS può inoltre usare questo messaggio per aggiornare i dati di una entità nel caso venga determinato un problema di sincronizzazione. Il messaggio contiene l'indirizzo TALQ di una delle entità per il quale il CMS richiede un aggiornamento. Può essere associato il parametro "since" per richiedere aggiornamenti a partire da un numero di sequenza particolare.

- *update<Entity>*: Messaggio di richiesta dal TALQ Brige al CMS per segnalare aggiornamenti di una entità. Questo messaggio è inviato dal TALQ Bridge ogni volta che c'è una modifica a carico di una entità. A messaggio ricevuto, il CMS risponderà con un HTTP OK. In assenza di OK, il TALQ Bridge ripeterà il messaggio fino a conferma avvenuta. Analogamente al messaggio di aggiornamento in direzione inversa, anche in questo caso è possibile effettuare un update parziale o totale dei dati. Il contenuto del messaggio è una entità individuale, i cui dati sono specifici rispetto all'entità e descritti in ogni messaggio inerentemente ai servizi interessati.

Il processo comunicativo viene rappresentato schematicamente in Figura.
Il possessore di una entità ha la responsabilità di mantenerne il numero di sequenza. Questo numero sarà incrementato con ogni aggiornamento dei dati associati all'entità. Il numero di sequenza sarà inviato insieme ai dati in ogni messaggio di tipo <*entity*> o *update<Entity>* usando l'attributo xml *"seq"*. Il ricevente del messaggio mantiene una copia dell'ultimo numero di sequenza ricevuto. Se un messaggio viene ricevuto con numero di sequenza superiore a quello in memoria (quindi cronologicamente successivo), il ricevente accetterà e gestirà il messaggio. Se è inferiore (fuori sequenza), ignorerà il messaggio. I numeri di sequenza permettono di sincronizzare in maniera efficiente CMS e TALQ Bridge, e gestire in maniera efficace messaggi duplicati e comunicazioni asincrone.
Per permettere un minor trasferimento di dati ridondanti, il protocollo prevede, per le richieste di aggiornamento (*refresh<Entity>* e *get<Entity>*) un attributo xml *"since"* dentro il quale il richiedente può specificare il suo numero di sequenza più recente per l'entità, e quindi, il limite inferiore richiesto per gli aggiornamenti. In assenza di questo attributo, viene inviato un aggiornamento completo.

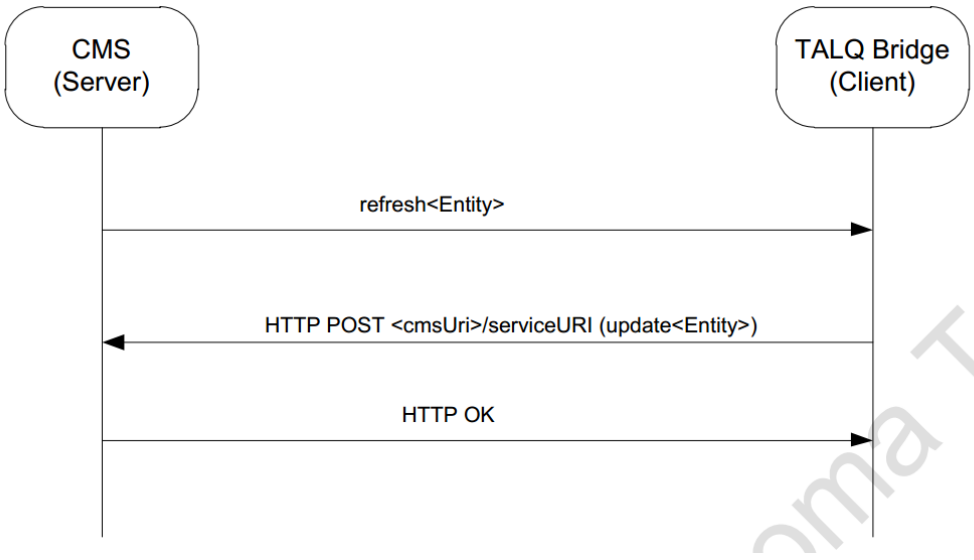


Figura 24 – Comunicazione tra TALQ Bridge e CMS per gestire una entità appartenente al TALQ Bridge. Il CMS richiede un aggiornamento dello stato della entità, il TALQ Bridge risponde con l'aggiornamento (totale o parziale), ad avvenuta ricezione il CMS conclude la comunicazione con HTTP OK.

I messaggi contenenti i dati degli aggiornamenti (i.e. *<entity>* ed *update<Entity>*), in risposta ai messaggi di richiesta aggiornamenti, sono dotati di un attributo xml *"modifier"* che può avere i seguenti valori:

- *complete:* I dati inviati sono una copia completa dei dati di entità. Il ricevente sostituirà tutti i dati in suo possesso per la specifica entità con quelli appena ricevuti.
- *update:* I dati rappresentano un aggiornamento parziale dell'entità, ed includono esclusivamente delle modifiche. Il ricevente modificherà esclusivamente i dati ricevuti, lasciando inalterata la rimanente parte dell'entità nella sua memoria.
- *delete:* L'entità non esiste più. Il ricevente cancellerà tutti i dati relativi alla entità.

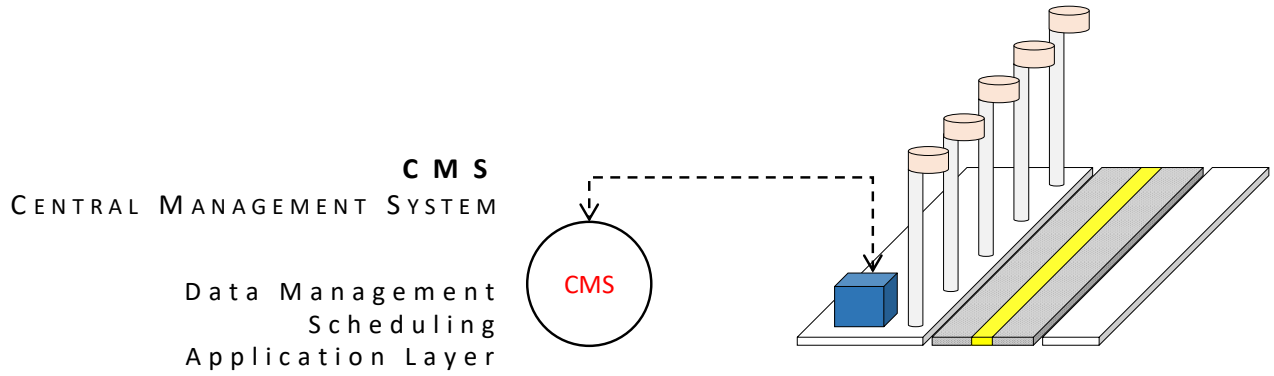
Nel caso venga ricevuto un messaggio di refresh<Entity> o get<Entity> per entità che non esistono, il ricevente assumerà che le entità siano state cancellate, e di conseguenza, invierà un messaggio <entity> o update<Entity> con attributo xml "modifier" impostato a *delete*. I messaggi riferiti alla cancellazione di un oggetto possono andare perduti, o ignorati a causa di mancata sincronia tra le parti. Sia il TALQ Bridge sia il CMS si re-sincronizzano quindi re-inviando un messaggio di *delete* in risposta.

2.3.11 Sincronizzazione dispositivi logici

La sincronizzazione tra i dispositivi logici segue una lista di regole che ne permette un controllo agevole essendo la *ownership* divisa tra CMS e TALQ Bridge.

1) Tutti gli attributi di configurazione saranno inizialmente posseduti dal lato creatore dell'entità, ovvero il TALQ Bridge.
2) Ciascun lato (CMS/TALQ Bridge) conserva due numeri di sequenza: uno per i propri attributi, e uno in copia del numero di sequenza dell'altro lato.
3) Il CMS può prendere possesso di un attributo inviando un aggiornamento su di esso al TALQ Bridge. Il TALQ Bridge non permetterà più modifiche locali all'attributo, e non aggiornerà più il proprio numero di sequenza al ricevimento di aggiornamenti da parte del CMS.
4) Il TALQ Bridge può richiedere nuovamente il possesso dell'attributo inviando un messaggio *takeAttributeOwnership* al CMS. Il passaggio di proprietà avviene solo dopo una risposta affermativa da parte del CMS con messaggio HTTP OK.
5) Il numero di sequenza del TALQ Bridge sarà usato come attributo "*since*" nei messaggi *refreshDeviceConfigurations* (i.e. messaggi per richiedere aggiornamenti) e come attributo "*seq"* nei messaggi *deviceConfigurations* e *deviceConfigurationsChanged* (i.e. messaggi per comunicare aggiornamenti).
6) Il numero di sequenza del CMS sarà usato come attributo *"since"* nei messaggi *getDeviceConfigurations* (i.e. messaggi per richiedere aggiornamenti) e come attributo *"seq"* nei messaggi *deviceConfigurations* e *deviceConfigurationsChanged* (i.e. messaggi per comunicare aggiornamenti).
7) Se il CMS riceve un aggiornamento per un attributo di cui ha preso possesso, o di cui sta attualmente prendendo possesso, senza aver ricevuto una esplicita richiesta dal TALQ Bridge di riprenderne possesso, il CMS ignorerà questo messaggio. Se il CMS ha già ricevuto conferma di passaggio di proprietà, sovrascriverà l'attributo con l'ultimo valore disponibile per riottenere il possesso. Questa regola permette di risolvere problemi di sincronizzazione che possono avvenire in caso i passaggi di proprietà richiedano tempo, e delle modifiche sugli attributi avvengano prima del completo svolgimento del passaggio.

2.3.12 Scenario Applicativo – Smart Street con Protocollo TALQ



SMART-LIGHT ACTUATOR (LOGIC DEVICE)
- Sistema di illuminazione modulabile (dimming)
- Sensore integrato di luminosità (misura alla sorgente)

PAVIMENTAZIONE A RECUPERO ENERGETICO
- Alimentazione locale network di sensori.

LIGHTING SMART-SENSOR (LOGIC DEVICE)
- Rilevazione illuminazione manto stradale (misura a terra)
- Rilevazione traffico

INTERFACCIA LOCALE (TALQ BRIDGE)
- Notification
- Configuration
- Lighting Control

Figura 25 – Caso di studio per illuminazione *smart* del manto stradale.

A seguito dello studio effettuato, è stato teorizzato uno scenario applicativo che può essere preso come caso di studio di fattibilità per il protocollo TALQ applicato all'illuminazione intelligente della pavimentazione stradale. Il sistema è rappresentato in Fig. 25. La rete di illuminazione esterna OLN è rappresentativo di una sezione di area stradale illuminata da dispositivi verticali. Gli attuatori principali preposti all'illuminazione sono realizzati con lampade ad intensità luminosa regolabile, dotati di sistemi di misura per potenza assorbita (individuale), inquinamento armonico in corrente (aggregato), fattore di potenza (aggregato), livello di luminosità (individuale). Per ogni palo (physical asset) sono presenti due dispositivi fisici, uno costituito da un array di attuatori luminosi, uno da un array di sensori elettrici e luminosi. A ciascun dispositivo fisico è associato un dispositivo logico con le funzioni necessarie.

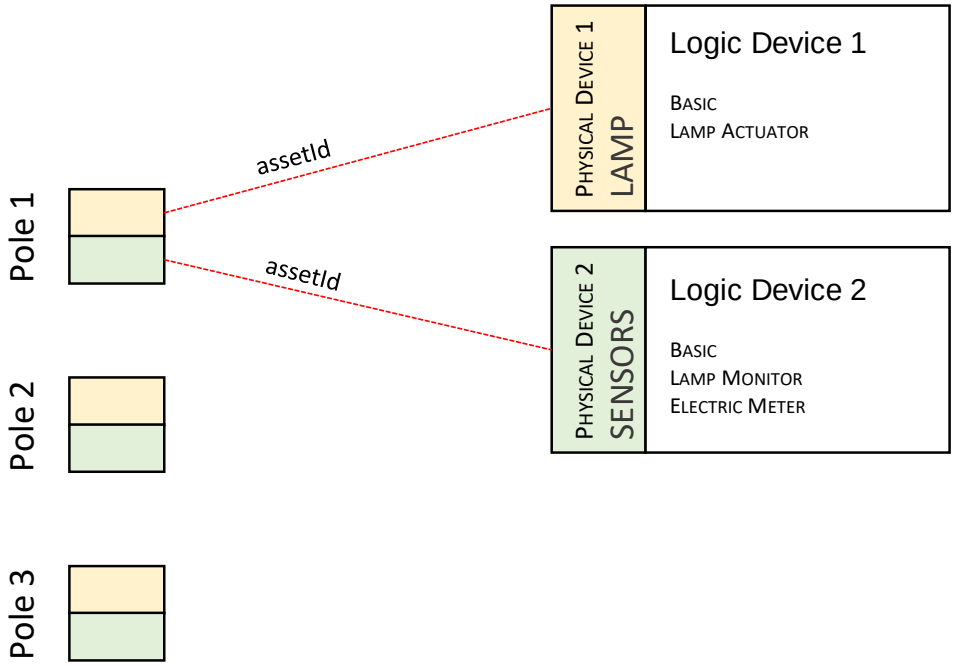


Figura 26 – Dispositivi fisici (Lampada e Sensore), dispositivi logici e funzioni implementate.

Un ulteriore sensore centralizzato di corrente e tensione è presente per la rilevazione locale del fattore di potenza e del livello di inquinamento armonico introdotto dai sistemi di illuminazione. Questo ultimo parametro tiene sotto controllo possibili problemi introdotti da carichi non lineari quali sono i sistemi di illuminazione a LED ad intensità regolabile. A livello di manto stradale sono inclusi dei sensori fotoelettrici per la rilevazione del livello di luminosità a terra, e dei sensori di tipo magnetico per la stima del traffico. L'alimentazione di questi sensori è demandata ad un sistema di *energy harvesting* implementato nel manto stradale stesso, per il recupero di energia elettrica dalle vibrazioni dei veicoli in passaggio. Le reti dei sensori a terra, degli attuatori luminosi e dei relativi sensori, sono connesse a delle centraline distribuite che svolgono il ruolo di TALQ Bridge. L'insieme delle centraline è connesso tramite rete wireless al CMS, il quale elabora le informazioni di consumo ed inquinamento energetico, di traffico automobilistico, di illuminazione e di condizioni atmosferiche, per determinare in tempo reale la configurazione ottimale per l'illuminazione locale. In caso di mancanza di comunicazione tra CMS e TALQ Bridge, le centraline dispongono di un programma elementare di funzionamento autonomo di tipo *day&night*.

2.3.13 Studio Protocollo TALQ - Conclusioni e sviluppi

È stato studiato un protocollo che descrive un sistema di gestione, basato su architettura Client-Server, per una rete di illuminazione di dimensioni estese. Sono state verificate le potenzialità del protocollo, sia in termini di flessibilità, sia in termini di praticità di implementazione. Il protocollo può essere usato come guida nella realizzazione dei sistemi di immagazzinamento dati, di messaggistica tra le parti interessate, e di controllo per i sensori e per gli attuatori che costituiscono le unità terminali della rete. Lo sviluppo degli applicativi adibiti alla gestione della comunicazione e della archiviazione dati in una rete di illuminazione ottemperante alle specifiche del protocollo TALQ è indubbiamente un compito complesso, ma permette di ottenere uno strumento in grado di interfacciarsi con uno standard industriale riconosciuto ed adottato a livello mondiale.

3 Conclusioni

In questo documento sono analizzate tre importanti problematiche relative alla innovazione tecnologica, funzionale e gestionale della illuminazione pubblica e degli ambienti confinati. Innanzitutto, i metodi e le tecnologie per la gestione e l'analisi di grandi quantità di dati di carattere energetico provenienti dai sistemi di illuminazione pubblica. È stato poi discusso il problema della cyber security, ovvero delle possibili vulnerabilità dei sistemi di illuminazione pubblica in ambito urbano. È stato infine illustrato un protocollo di comunicazione standard per la trasmissione di informazioni riguardanti l'illuminazione pubblica.

4 Riferimenti bibliografici

1. Bari, Ataul, et al. "Challenges in the smart grid applications: an overview. "*International Journal of Distributed Sensor Networks* 2014 (2014).
2. Report to NIST on Smart Grid Interoperability Standards Roadmap EPRI, Jun. 17, 2009 [Online].
3. K. Moslehi and R. Kumar, "Smart grid—A reliability perspective," Innovative Smart Grid Technologies (ISGT), pp. 1–8, Jan. 19–21, 2010.
4. Vehbi C. Güngör, Member, IEEE, Dilan Sahin, Taskin Kocak, Salih Ergüt, Concettina Buccella, Senior Member, IEEE, Carlo Cecati, Fellow, IEEE, and Gerhard P. Hancke, Senior Member, IEEE. Smart Grid Technologies: Communication Technologies and Standards.
5. S. Zeadally, A. Pathan, C. Alcaraz, and M. Badra, "Towards Privacy Protection in Smart Grid", Wireless Personal Communications, vol. 73, pp. 23- 50, 2012.
6. Ye Yan, Yi Qian, Hamid Sharif and David Tipper "A Survey on Cyber Security for Smart Grid Communications".
7. Q.Yang, J. "Afor distributed control of power distribution networks," IEEE Trans. Barria, and T. C. Green, "Communication infrastructures Ind. Inform., vol. 7, no. 2, pp. 316–327, May 2011.
8. Hrasnica, Halid, Abdelfatteh Haidine, and Ralf Lehnert. *Broadband powerline communications: network design*. John Wiley & Sons, 2005.
9. K. Dostert. *Telecommunications over the Power Distribution Grid – Possibilities and Limitations*. IIR-Powerline 6/97, Germany, 1997.
10. Zimmermann, Hubert. "OSI reference model-the ISO model of architecture for open systems interconnection." *IEEE Transactions on communications* 28.4 (1980): 425-432.
11. Yigit, Melike, et al. "Power line communication technologies for smart grid applications: a review of advances and challenges." Computer Networks 70 (2014): 366-383.
12. Lee, M. K., et al. "HomePlug 1.0 powerline communication LANs—protocol description and performance results." International Journal of Communication Systems 16.5 (2003): 447-473.
13. Lee, Jae-Min, et al. "A new home network protocol for controlling and monitoring home appliances-HNCP." Interantional conference on consumer electronics. 2002.
14. TALQ Technical Working Group, "TALQ Specification Overview". White paper, March 2015. www.talq-consortium.org
15. Gungor, Vehbi C., and Frank C. Lambert. "A survey on communication networks for electric system automation." Computer Networks 50.7 (2006): 877-897.
16. Kabalci, Yasin. "A survey on smart metering and smart grid communication. "Renewable and Sustainable Energy Reviews 57 (2016): 302-318.
17. Yan, Ye, et al. "A survey on cyber security for smart grid communications."IEEE Communications Surveys & Tutorials 14.4 (2012): 998-1010.
18. Liu, Yao, Peng Ning, and Michael K. Reiter. "False data injection attacks against state estimation in electric power grids." ACM Transactions on Information and System Security (TISSEC) 14.1 (2011): 13.
19. McDaniel, Patrick, and Stephen McLaughlin. "Security and privacy challenges in the smart grid." IEEE Security and Privacy 7.3 (2009): 75-77.

20. Falliere, Nicolas, Liam O. Murchu, and Eric Chien. "W32. stuxnet dossier."White paper, Symantec Corp., Security Response 5 (2011): 6.
21. Wang, Wenye, and Zhuo Lu. "Cyber security in the Smart Grid: Survey and challenges." Computer Networks 57.5 (2013): 1344-1371.
22. Mirkovic, Jelena, and Peter Reiher. "A taxonomy of DDoS attack and DDoS defense mechanisms." ACM SIGCOMM Computer Communication Review34.2 (2004): 39-53.
23. [Biba, Kenneth J. Integrity considerations for secure computer systems. No. MTR-3153-REV-1. MITRE CORP BEDFORD MA, 1977.
24. Fraser, Timothy. "LOMAC: Low water-mark integrity protection for COTS environments." Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000.
25. Clark, David D., and David R. Wilson. "A comparison of commercial and military computer security policies." Security and Privacy, 1987 IEEE Symposium on. IEEE, 1987.
26. Efthymiou, Costas, and Georgios Kalogridis. "Smart grid privacy via anonymization of smart metering data." Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on. IEEE, 2010.
27. Kim, Sungwook, et al. "A secure smart-metering protocol over power-line communication." IEEE Transactions on Power Delivery 26.4 (2011): 2370-2379.
28. Gupta, B. B., et al. "Fighting against phishing attacks: state of the art and future challenges." Neural Computing and Applications (2016): 1-26.
29. Ollmann, Gunter. "The Phishing Guide–Understanding & Preventing Phishing Attacks." NGS Software Insight Security Research (2004).
30. Anti-Phishing Working Group (APWG), ‘‘Phishing activity trends report - first quarter 2013.
31. Aloul, Fadi A. "The need for effective information security awareness." Journal of Advances in Information Technology 3.3 (2012): 176-183.
32. James L (2005) Phishing exposed. Syngress Publishing, Burlington.
33. Jakobsson, Markus, and Steven Myers, eds. Phishing and countermeasures: understanding the increasing problem of electronic identity theft. John Wiley & Sons, 2006.
34. Jagatic, Tom N., et al. "Social phishing." Communications of the ACM 50.10 (2007): 94-100.
35. Granger, Sarah. "Social engineering fundamentals, part I: hacker tactics." Security Focus, December 18 (2001).
36. Jagatic, Tom N., et al. "Social phishing." Communications of the ACM 50.10 (2007): 94-100.
37. Spear Phishing Attacks - Why They are Successful and How to Stop Them. Combating the Attack of Choice for Cyber criminals, Fire Eye Inc (Whitepaper).
38. Chou, Neil, et al. "Client-Side Defense Against Web-Based Identity Theft." NDSS. 2004.
39. Downs, Julie S., Mandy Holbrook, and Lorrie Faith Cranor. "Behavioral response to phishing risk." Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit. ACM, 2007.
40. Huang H, Tan J, Liu L (2009) Countermeasure techniques for deceptive phishing attack. In: International conference on new trends in information and service science, 2009. NISS’09, Korea, pp 636–641.
41. Sheng, Steve, et al. "Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2010.
42. Dong, Xun, John A. Clark, and Jeremy Jacob. "Modelling user-phishing interaction." 2008 conference on human system interactions. IEEE, 2008.
43. Kumaraguru, Ponnurangam, et al. "Protecting people from phishing: the design and evaluation of an embedded training email system." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 2007.
44. Levine J (2008) DNS blacklists and whitelists, IRTF anti-spam research group, Nov 2008.
45. Microsoft, Sender ID, 2008. http://www.microsoft.com/. Accessed on Sept 2014
46. Sheng, Steve, et al. "An empirical analysis of phishing blacklists." Proceedings of Sixth Conference on Email and Anti-Spam (CEAS). 2009.

47. Nathan Marz and James Warren. “Big Data: Principles and best practices of scalable realtime data systems”. Manning Publications, 2015.
48. Apache Hadoop, http://hadoop.apache.org/
49. Apache Spark, http://spark.apache.org/
50. Galli, B. S., Ieee, S. M., Scaglione, A., & Ieee, F. (2011). “For the Grid and Through the Grid : The Role of Power Line Communications in the Smart Grid”. Proceedings of the IEEE, 99(6).
51. Ataul Bari, Jin Jiang, Walid Saad, and Arunita Jaekel, “Challenges in the Smart Grid Applications: An Overview,” International Journal of Distributed Sensor Networks, vol. 2014, Article ID 974682, 11 pages, 2014. doi:10.1155/2014/974682
52. Barenghi, Alessandro, et al. "Computer Security Anchors in Smart Grids: The Smart Metering Scenario and Challenges." Trusted Computing for Embedded Systems. Springer International Publishing, 2015. 47-59.
53. Panzieri, Stefano, Ilaria Scarano, and Roberto Setola. "Vulnerabilità informatica dei sistemi SCADA connessi alle reti pubbliche." Valutazione e gestione del rischio negli insediamenti civili ed industriali. VGR 4 (2004).
54. PRIME Specification, http://www.prime‐alliance.org/
55. PRIME Technology Whitepaper. PHY, MAC, and Convergence layers. PRIME Project, July 2008. http://www.prime‐alliance.org/wp-content/uploads/2013/03/MAC_Spec_white_paper_1_0_080721.pdf
56. PRIME v1.4 Evolution: A Future Proof of Reality Beyond Metering, A. Sendin, I. H. Kim, S. Bois, A. Munoz, A. Llano, 5th IEEE International Conference on Smart Grid Communications (SmartGridComm), Nov. 2014.
57. G3-PLC, http://www.g3-plc.com/
58. HomePlug “Green PHY™ 1.1 - The Standard for In-Home Smart Grid - Powerline Communications: An application and technology overview”, HomePlug Powerline Alliance, Inc. All rights reserved. Version 1.02, 3 October 2012
59. TAQL consortium, http://www.talq-consortium.org

5 Curriculum scientifico del gruppo di lavoro

Stefano Panzieri è Professore Associato di Automatica presso Roma TRE. Vicepresidente del Comitato Unico di Garanzia dell'Ateneo Roma TRE. Responsabile del Laboratorio di Modellistica e Simulazione nel settore della Protezione delle Infrastrutture Critiche (MCIP Lab). Coordinatore di Ateneo del progetto per la Regione Lazio Smart Environments. Responsabile di ricerche sul tema della diagnostica energetica nell'ambito Smart Buildings. Coordinatore di alcuni progetti Europei sulle Infrastrutture Critiche. Coordinatore del Dottorato di Ricerca in Ingegneria Informatica e dell'Automazione. Dottore di Ricerca in Ingegneria dei Sistemi nel 1993 alla Sapienza. La sua attività di insegnamento si svolge nel settore dei Controlli Automatici, nei Sistemi di Controllo di Processo e nella gestione della sicurezza di grandi infrastrutture. Autore di più di 150 pubblicazioni in ambito internazionale sulle tematiche già citate e nel settore della Robotica

Riccardo Torlone è professore ordinario di ingegneria informatica presso Roma Tre. È Coordinatore del gruppo di ricerca su big data e basi di dati, Coordinatore del Collegio Didattico di Ingegneria Informatica e membro della Giunta del Dipartimento di Ingegneria dell'Università Roma Tre. È stato visiting researcher presso la University of California Los Angeles (UCLA). La sua attività di ricerca ha riguardato vari argomenti nel settore delle basi di dati e dei sistemi informativi. È autore di più di 150 articoli pubblicati sulle principali riviste del settore e negli atti delle principali conferenze. È coautore del libro di testo sulle basi di dati più diffuso in Italia, pubblicato in sei edizioni e articolato in due volumi. È responsabile di progetti di ricerca

finanziati dalla Commissione Europea e da numerosi enti pubblici e privati. Insegna un corso di Big Data presso Roma Tre.

Federica Pascucci si è laureata in Ingegneria Informatica presso l'Università degli Studi Roma Tre nel 2000 e ha conseguito il titolo di dottore di ricerca in Ingegneria dei Sistemi nel 2004 presso l'Università degli Studi di Roma "La Sapienza". I suoi interessi di ricerca includono la localizzazione indoor, sistemi di posizionamento per treni, le reti di sensori, la diagnosi di guasti cyber e fisici in impianti industriali e infrastrutture critiche. Ha pubblicato più di 70 paper su riviste e in atti di convegni internazionali ricevendo diversi riconoscimenti internazionali. E' stata responsabile di unità di ricerca in diversi progetti nazionali e internazionali. E' docente dei corsi di "Controllo digitale", "Identificazione e fusione sensoriale" e "Cyber Physical Systems" presso l'Università degli Studi Roma Tre.

Giuseppe Bernieri ha conseguito la laurea magistrale nel luglio 2014 in Ingegneria Gestionale e dell'Automazione presso l'Università degli Studi Roma Tre e dal novembre dello stesso anno è studente di Dottorato del Dipartimento di Ingegneria della stessa Università. I suoi interessi di ricerca riguardano la sicurezza cyber applicata agli impianti industriali e alle infrastrutture critiche, lo sviluppo e l'implementazione di soluzione ad-hoc per sistemi di controllo e sistemi SCADA. Collabora attivamente con il Network, Information and Security Laboratory del dipartimento di informatica dell'Università di Malaga. Ha partecipato a diversi progetti europei come FACIES e RISING. E' tutor dei corsi di "Controllo digitale" e "Cyber Physical Systems" presso l'Università degli Studi Roma Tre.