Information Fusion e sorveglianza

F. Flammini, A. Pappalardo, V. Vittorini

Uno degli ambiti maggiormente interessati dall'aumento del rapporto tra numero di elaboratori e numero di utenti è quello dei sistemi di monitoraggio distribuito per applicazioni di supervisione, diagnostica e sorveglianza. In tali sistemi, infatti, pochi operatori sono chiamati a controllare un elevato numero di sensori e sottosistemi, spesso distribuiti ed eterogenei. In questo articolo verranno illustrate tecniche automatiche di integrazione ed interpretazione dei dati per consentire di gestire una mole potenzialmente elevata di informazioni ed adottare decisioni efficaci in tempi rapidi.

Keywords: *Physical Security, Smart-sensing, Survelliance, Information Fusion*

1. Introduzione

Le reti di sensori costituite da dispositivi eterogenei sono adottate in molti ambiti al fine di rilevare eventi di diversa natura e supportare di conseguenza le decisioni degli operatori delle sale di controllo. Si pensi ad esempio ai seguenti contesti:

- monitoraggio ambientale per l'allerta precoce (*early warning*) di eventi potenzialmente disastrosi quali terremoti, alluvioni, smottamenti, incendi, etc.
- sistemi di allarme, sicurezza fisica[1] e sorveglianza per il rilevamento di intrusioni (*intrusion detection*) in aree protette, comportamenti anomali, individui sospetti, sostanze pericolose, etc.

[1] *I sistemi di supervisione applicati alla sicurezza fisica sono stati recentemente battezzati come PSIM (Physical Security Information Management), per specializzarli rispetto ai più tradizionali SCADA (Supervisory Control and Data Acquisition).*

- prognostica e diagnostica, per anticipare, rilevare e gestire guasti o incidenti in centrali, impianti industriali di produzione e in altre infrastrutture critiche.

In questi ed altri contesti emergenti, per ragioni di economicità e facilità di installazione, i sensori tradizionali sono sempre più affiancati o soppiantati da moderni *smart-sensors*, quali i cosiddetti *motes* delle Wireless Sensor Networks (WSN) [5], e *smart-cameras*, ovvero videocamere di rete "intelligenti", con un sistema di elaborazione dedicato a bordo. Tali dispositivi sono in grado di pre-elaborare le misure di parametri ambientali al fine di generare eventi anziché dati grezzi (*raw data*). Ad esempio, i sensori WSN possono notificare in automatico eventi quali temperatura, umidità o accelerazione oltre i limiti di normalità, mentre le telecamere intelligenti sono in grado di far funzionare sofisticati algoritmi di visione artificiale, per il rilevamento di presenza persone, attraversamento perimetri critici, oggetti incustoditi, elevato affollamento ed altro ancora [8].

La rapida crescita ed evoluzione dei sistemi di monitoraggio distribuito sopracitati fa sì che il numero di sensori, e quindi la mole di eventi potenzialmente generati (incluse le segnalazioni non attendibili, ovvero i cosiddetti "falsi positivi"), non sia compatibile con la capacità di gestione in tempo reale di un numero ridotto di operatori umani. Quello che deve essere realizzato, quindi, è un compromesso tra l'opportunità di raccogliere un numero di informazioni il più elevato possibile, per incrementare le capacità di riconoscimento e consapevolezza situazionale (brutta versione italiana dell'inglese *situation awareness*), ed il rispetto dei requisiti di usabilità ed ergonomia (i cosiddetti *human factors*), che impattano sulla correttezza e sulla tempistica della risposta agli eventi. E' quindi una delle molteplici applicazioni, in un settore più che mai interdisciplinare, dell'ambito di ricerca comunemente noto come *information fusion* per il supporto alle decisioni (si veda il Riquadro 1).

Da quanto detto, dovrebbe essere chiaro che i sistemi intelligenti di monitoraggio distribuito impiegati in attività di supervisione, diagnostica e sorveglianza, necessitano di un continuo processo di innovazione, da un lato per soddisfare requisiti applicativi in rapida evoluzione, dall'altro per affrontare le criticità ed i problemi aperti che si vengono a creare come naturale conseguenza di tale evoluzione. Gli aspetti più coinvolti da questo processo riguardano in particolare:

- La problematica generale della gestione di dati provenienti da molte sorgenti eterogenee di informazione da parte di un ridotto numero di utenti [7].
- La specializzazione al caso dei sistemi di supervisione, monitoraggio distribuito e sorveglianza, in cui un limitato numero di operatori deve prendere decisioni in funzione di segnalazioni provenienti da un numero elevato di sottosistemi e apparati sensoriali [13].

- La disponibilità di sistemi di integrazione e *reasoning* di tali informazioni allo scopo di supportare le decisioni degli operatori [12].

Questi sono gli aspetti che andremo ad affrontare in questo articolo, partendo dalle esigenze industriali che emergono da scenari applicativi reali (descritti nel paragrafo 2) ed illustrando i diversi ambiti di ricerca coinvolti, ivi comprese le tecnologie di rilevamento multi-modale e multi-mediale e le tecniche di aggregazione delle informazioni (paragrafi 3-6), escludendo dal nostro ambito di analisi gli aspetti più propriamente *human-related* relativi alla percezione e alla cognizione, a cui si daranno dei rapidi cenni e qualche puntatore alla letteratura di riferimento.

2. Information Fusion applicata alla sorveglianza

Il progresso delle tecnologie di sorveglianza registrato negli ultimi anni ha permesso di soddisfare requisiti applicativi in rapida evoluzione, ma ha anche creato aspettative irrealistiche da parte degli utenti, spesso poco consapevoli del rischio non trascurabile di falsi allarmi e dei loro effetti collaterali. Un'altra considerazione è che ha poco senso disporre di tecnologie eterogenee, come richiesto in ambienti complessi, senza sfruttarle adeguatamente. Pertanto, è evidente la necessità di comprendere come le tecnologie possano essere combinate ed integrate, per trovare il miglior compromesso tra costi, complessità di gestione e benefici attesi di un sistema di monitoraggio in un determinato ambito applicativo.

Tali aspetti sono strettamente connessi alle capacità di supporto alle decisioni che ci aspetta dai sistemi di supervisione, diagnostica e sorveglianza. In questo senso, supportare in maniera affidabile il processo di *decision-making* è fondamentale per evitare interpretazioni errate delle informazioni che potrebbero condurre all'attivazione di contromisure non appropriate.

Metodi, tecniche e strumenti disponibili a tal scopo dovrebbero essere inquadrati in un paradigma che tenga conto dei componenti principali del sistema e del ruolo fondamentale dell'operatore umano. Da una parte, c'è la necessità di inserire la prospettiva dell'utente nella progettazione di un sistema di *Information Fusion* (IF) [3]; dall'altra c'è bisogno di impiegare l'IF nella progettazione di un sistema di *Decision Support* (DS) per migliorare l'intero processo decisionale [12]. In altre parole, è necessaria la convergenza tra queste due aree di ricerca che, sebbene intimamente connesse, si sono sviluppate negli anni in maniera separata.

Più in dettaglio, scenari applicativi reali evidenziano esigenze che si pongono su livelli differenti: capacità funzionali, architetture di riferimento e grado di interazione con gli utenti.

2.1. Livelli di information fusion

Per prima cosa, bisogna definire il livello di astrazione della strategia di fusion [1]. In altre parole, è necessario comprendere se gli input del processo sono dati grezzi misurati dai sensori o dati pre-elaborati con l'intento di estrarre già un'informazione di più alto livello semantico. Analogamente, va definito il tipo di output del processo, che rappresenta il fine ultimo per il quale viene implementata la IF. Dal punto di vista dell'operatore, la strategia di fusion è quella di alto livello (*Decision Fusion*). La motivazione risiede nel fatto che in questo modo è possibile ragionare a partire da informazioni allo stesso livello semantico, ottenuto omogeneizzando molteplici *input*, ognuno caratterizzato da una diversa rappresentazione delle informazioni. Inoltre, tale strategia di fusion offre una scalabilità che è difficile da ottenere con una IF di basso livello, il che è un requisito importante laddove le aree da monitorare sono potenzialmente estese. Ciò non vieta di adottare strategie multi-livello: a basso livello si impiegano metodi di fusion mono o multi-modale (come sarà meglio spiegato in seguito), mentre ad alto livello si impiegano dati già elaborati. A livello di output, la strategia da perseguire nel contesto in esame dovrebbe consentire di valutare aspetti diversi: i) stati e attributi delle entità monitorate; ii) relazioni reciproche tra le entità e con l'ambiente circostante; iii) stati futuri e proiezioni delle situazioni interpretate, per valutare minacce, rischi e possibili impatti.

2.2. Gerarchia di centri di controllo e operatori

Riconoscere le situazioni in corso di svolgimento, la loro fase di evoluzione e possibili trend, è funzionale all'adozione delle misure più appropriate in risposta a quanto rilevato. Un processo congiunto di IF e DS consente dunque di potenziare in maniera significativa le capacità degli operatori umani, che dai centri di controllo supervisionano gli apparati installati in campo ed eseguono le procedure di gestione degli eventi. Le azioni incluse in queste procedure sono spesso orchestrate dagli stessi sistemi di monitoraggio e sorveglianza, la cui architettura può essere distribuita e gerarchica, con un numero ridotto di sale di controllo centrali e molte più sale locali, che raccolgono le segnalazioni in base a scopi e responsabilità differenti [2]. Attraverso apposite interfacce, un operatore può supervisionare, prendere in carico e confermare l'esecuzione di ciascuna fase procedurale che è nella sua responsabilità. Integrando queste funzionalità con quelle più avanzate che conseguono dall'IF, è possibile reagire in maniera più tempestiva e puntuale alle situazioni di emergenza. Ad esempio, in caso di sistemi di supervisione distribuiti e gerarchici, è possibile riconoscere fenomeni complessi, quali attacchi terroristici di tipo strategico, che prevedono l'esecuzione contemporanea di azioni dolose nei confronti di *asset* distanti dell'infrastruttura monitorata [9].

2.3. Information fusion e fattori umani

Molte infrastrutture sono geograficamente estese e richiedono la protezione di molteplici risorse e tipologie di utenti (operatori, manutentori, pubblico). Per garantire una copertura adeguata di tali infrastrutture è necessario impiegare un numero elevato di dispositivi (telecamere, sensori anti-intrusione, rilevatori di sostanze chimiche, etc.). Ciò rende impraticabile una sorveglianza unicamente *human-based*, dato che l'analisi manuale di numerosi flussi video ed altre segnalazioni di allarme è molto impegnativa, soggetta a cali di attenzione ed errori di gestione. Dal momento che il numero di operatori umani è tipicamente limitato, l'analisi del fattore umano assume una notevole importanza [12] [6]. Da una parte, infatti, vi è la tendenza dei singoli sottosistemi integrati a produrre una mole elevata di dati (eventi, allarmi, messaggi diagnostici, etc.), dall'altra le ridotte capacità umane nel gestirli in tempo reale, specialmente nel caso di eventi critici simultanei. Inoltre, attività fortemente non stimolanti e ripetitive rendono all'atto pratico poco efficace la sorveglianza assistita dall'uomo [13]. Tecnologie come l'analisi video intelligente attenuano considerevolmente i problemi dovuti ai fattori umani, che attualmente rappresentano uno dei principali handicap nell'analisi dei dati relativi ad attività di sorveglianza. Tuttavia, come anticipato, ciò non è ancora sufficiente, in quanto le tecnologie disponibili allo stato dell'arte non hanno livelli di affidabilità elevatissimi oppure generano eventi che devono essere comunque inquadrati nel contesto, interpretati e verificati prima di adottare eventuali contromisure. Pertanto, il grado di dipendenza dal livello di attenzione degli operatori è ancora molto elevato.

Relativamente alle azioni di risposta agli eventi, il controllo dei falsi allarmi è fondamentale. Da questo punto di vista, tutti i sistemi automatici di diagnostica (o similari) dovrebbero garantire una soglia di affidabilità minima, al di sotto della quale il sistema diventa all'atto pratico inefficace: nello specifico, in base a recenti studi sull'ergonomia, tale soglia viene superata quando oltre il 30% delle segnalazioni prodotte dal sistema sono da scartare perché false o irrilevanti [14]. La riduzione del numero di falsi allarmi riduce i costi di gestione ed ha un significativo impatto sulle prestazioni degli operatori, influenzandone l'efficienza in termini di rapidità di risposta e confidenza nell'accuratezza del sistema. L'implementazione di un processo di IF applicato alla sorveglianza è dunque funzionale anche al miglioramento dell'affidabilità delle segnalazioni, che dipende sostanzialmente da due parametri: la probabilità di rilevamento (*Probability Of Detection*, indicata con POD) ed il tasso di falsi allarmi (*False Alarm Rate*, indicato con FAR).

3. Smart-sensing

Per raggiungere il livello richiesto di conoscenza relativo ad un'area da monitorare è necessaria la "cattura" di flussi informativi appropriati. Un sistema di sorveglianza efficace utilizza sensori per monitorare parametri ambientali, determinare il verificarsi di condizioni anomale, rilevare l'occorrenza di eventi indesiderati. Essi rappresentano dunque le principali sorgenti di informazione, che concorrono a raggiungere una opportuna situation awareness.

I dispositivi sensoriali sono basati su tecnologie più o meno sofisticate a seconda delle caratteristiche richieste in termini di funzionalità, prestazioni, costi, adattabilità a particolari condizioni ambientali, etc. Uno degli ambiti applicativi in cui si richiede che i sensori esibiscano sofisticate funzionalità ed alte prestazioni è, ad esempio, la protezione delle infrastrutture critiche in contesti di homeland security.

Le attuali esigenze di sorveglianza richiedono che i moderni sistemi non rivestano più solo un ruolo di supporto, ma diventino il cuore del processo di inferenza. Poiché l'"intelligenza" del sistema risiede anche nella capacità di sensing, si è assistito negli ultimi anni allo sviluppo e all'utilizzo sempre crescente dei cosiddetti sensori smart.

Per capire cosa si intenda per smart-sensor, è opportuno specificare cosa intendiamo per sensore "standard". In generale, un sensore è un dispositivo realizzato con l'obiettivo di effettuare una misura di una grandezza fisica relativa ad un oggetto o all'ambiente, e trasformarla in un segnale elettrico analogico o digitale. In Figura 1 è mostrata l'architettura logica di un sensore, che comprende le seguenti componenti principali: sensing unit (per l'acquisizione del segnale), processing unit (ad esempio per l'amplificazione, la compensazione e il filtraggio del segnale), power unit (preposta all'alimentazione elettrica), sensor interface (per il collegamento fisico del dispositivo con altre componenti elettroniche).

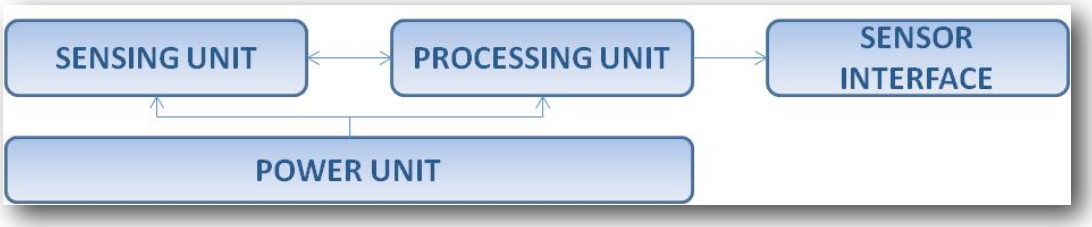

Figura 1
Architettura logica di un sensore

La principale differenza tra sensori “smart” e sensori “standard” risiede nel fatto che i primi ospitano a bordo sistemi *embedded* completi, dotati di microprocessore, unità di memoria, trasmettitori radio, e relativi software di base per consentirne la programmazione. Di tali sensori “evoluti” si sono date negli ultimi anni diverse definizioni, tra le quali le due più comunemente accettate ne sottolineano rispettivamente gli aspetti tecnologici (sensori “smart”) oppure le caratteristiche funzionali (sensori “intelligenti”). La distinzione si riferisce principalmente al ruolo assunto dal microprocessore: se è utilizzato ad esempio per l’elaborazione digitale dei segnali, il calcolo e le funzioni di interfacciamento, o se viene anche impiegato a supporto di funzionalità avanzate (“intelligenti”) quali l’auto-diagnostica, l’auto-adattamento, l’auto-identificazione, l’auto-calibrazione, etc. In alcuni casi, il microprocessore può essere utilizzato anche per decidere quando memorizzare dati o per minimizzare il consumo di energia.

Traendo vantaggio dei recenti sviluppi della tecnologia nell’ambito della microelettronica e dei sistemi micro elettro-meccanici (MEMS), la dimensione di tali sensori è andata riducendosi nel tempo, mentre la produzione intensiva per una grande varietà di applicazioni basata su tecnologia VLSI (*Very Large Scale Integration*) ha portato ad una forte riduzione del loro costo. Questi fattori hanno contribuito a rendere la produzione dei sensori sempre più conveniente, a fronte di un’”intelligenza” sempre maggiore. Nel prossimo futuro, il progresso tecnologico porterà un miglioramento anche nelle altre componenti (memoria, trasmissione radio, batterie) consentendo autonomia elevata e costi di manutenzione ridotti.

Le prime applicazioni di sensori intelligenti sono state attuate nell’ambito dell’ingegneria civile, ad es. nel monitoraggio strutturale (*Structural Health Monitoring*, SHM). Oggi, l’adozione di sensori evoluti è sempre più frequente in diversi ambiti. In particolare, è in aumento il loro utilizzo nell’ambito della difesa del territorio al fine di implementare strategie di allerta precoce (*early warning*). In tale contesto, la sorveglianza di un’area da proteggere si è spostata dalla tradizionale TVCC (televisione a circuito chiuso) a sistemi che integrano dispositivi molto eterogenei e con differenti livelli di intelligenza, classificabili in determinate categorie funzionali [11]. Tra i sensori più noti rientrano quelli per: il rilevamento delle intrusioni in aree protette tramite radar ad infrarosso, microonde o ultrasuoni, il riconoscimento tramite *smart-cameras* di comportamenti anomali (es. bagaglio incustodito); il monitoraggio di parametri ambientali quali temperatura, luminosità, umidità, accelerazione, tramite *motes*; l’analisi audio intelligente per rilevare urla, spari e rottura vetri; il rilevamento di agenti nocivi (CBRNe: Chimico, Biologico, Radiologico, Nucleare, esplosivo); l’ispezione a distanza di una massa di individui per rilevare armi nascoste, tramite onde millimetriche a terahertz, etc.

4. Sorveglianza multi-modale e multi-mediale

Sebbene la video sorveglianza rappresenti la forma più popolare di monitoraggio, ne esistono diverse altre in via di diffusione. Come evidenziato nella precedente sezione, l'impiego di nuovi dispositivi è da un lato incoraggiato dagli avanzamenti tecnologici, dall'altro richiesto da requisiti in continua evoluzione. L'integrazione di tecnologie eterogenee conduce a sistemi di sorveglianza multi-modali e multi-mediali [14], in cui la diversità tecnologica rappresenta un efficace valore aggiunto. I sistemi di sorveglianza multi-mediali collezionano ed elaborano diversi flussi di informazioni: video, audio, e qualunque altro output provenga da sensori installati nella stessa area. Pertanto si può affermare che un sistema di sorveglianza multi-modale e multi-mediale (ovvero capace di fornire informazioni di diversa natura, catturate attraverso mezzi differenti) combina due caratteristiche principali:

- l'utilizzo di più sensori connessi in rete, con la possibilità di sovrapporre parzialmente le aree di monitoraggio;
- l'utilizzo di sensori eterogenei, in modo da rilevare più informazioni di interesse disponibili all'interno dell'area monitorata.

La combinazione può essere realizzata adottando diverse soluzioni. Ad esempio, una configurazione elementare può essere basata su videocamere intelligenti o su sensori acustici, appartenenti a classi diverse e quindi con diverse caratteristiche e potenzialità. La soluzione basata su videocamere potrebbe ad esempio comprendere dispositivi brandeggiabili (PTZ: *Pan-Tilt-Zoom*), con un campo visivo modificabile e regolabile da remoto, oltre che videocamere fisse, e con tecnologie ad infrarossi, termiche, etc. Questo tipo di soluzione è generalmente adottata per migliorare ed estendere la visione della scena sfruttando efficacemente *ridondanza* e *sovrapposizione* nell'installazione dei dispositivi. Da un punto di vista prettamente funzionale, una simile rete di video camere consente di catturare maggiori dettagli relativi agli eventi identificati e di coprire aree estese; ad esempio è possibile effettuare automaticamente lo zoom di un intruso e identificarlo, tracciare oggetti o persone sospette, ed effettuare l'analisi di una folla di persone. Inoltre, specifiche combinazioni di videocamere fisse e PTZ possono essere utilizzate, attraverso un approccio distribuito, regolando in automatico alcuni parametri dei dispositivi installati (come la risoluzione) a seconda della complessità della scena osservata (ad esempio in presenza di un numero più o meno alto di persone). In questo modo una PTZ può osservare una singola persona con alta risoluzione, o tracciare più persone allo stesso tempo ad una risoluzione inferiore.

Più in generale, differenti punti di vista di una stessa scena possono risolvere eventuali problemi dovuti ad occlusioni, guasti e sabotaggi. Questo è un punto

particolarmente importante nel caso di una videosorveglianza intelligente, ovvero realizzata con il supporto di una video-analisi via software dei flussi video catturati dalle telecamere. In tal caso infatti, le prestazioni di un algoritmo di *video analytics* dipendono anche dalle dimensioni dell'area sottoposta a sorveglianza, all'interno della quale il numero di elementi occlusivi (persone, pilastri, muri, o altri oggetti che ostruiscono il campo visivo) hanno evidentemente un impatto sul rilevamento ed il tracciamento di un obiettivo.

E' possibile pertanto migliorare la qualità del monitoraggio, e di conseguenza superare alcuni limiti degli algoritmi di video analisi anche in ambienti non favorevoli ed in condizioni di affollamento. Questi benefici hanno alcuni contro-altari, quali la difficoltà che si può incontrare nella corretta calibrazione dei dispositivi e nella corretta gestione delle videocamere disponibili. Quest'ultimo elemento è peraltro fondamentale per l'espletamento di funzionalità quali il tracciamento, essendo di grande importanza la corretta gestione dell'*hand-over* tra una video camera e l'altra, in modo da non perdere l'oggetto tracciato e la sua identità durante il suo spostamento.

Oltre alle occlusioni, l'efficacia della video-analisi può risentire anche di improvvisi cambiamenti di luminosità, della presenza di superfici riflettenti o di oggetti che si confondono con lo sfondo (*background*) della scena. In generale, la sorveglianza in situazioni complesse non può essere condotta soltanto attraverso video camere, ma è necessario – come già menzionato – acquisire ed integrare diverse fonti di informazione.

A questo va aggiunto che, per soddisfare le richieste sempre più sofisticate degli utenti finali, la complessità degli eventi che si desidera rilevare è sempre maggiore. Tipicamente, l'approccio alla base di queste configurazioni multimodali più avanzate può essere di tipo *bottom-up* o *top-down* .

L'approccio *bottom-up* è utilizzato per realizzare integrazioni ad-hoc, in cui i sensori sono considerati come *black-box* e non utilizzano nelle loro elaborazioni le informazioni provenienti da altri sensori. La logica di integrazione, quindi, è esterna e risiede ad un più alto livello di astrazione.

Nell'approccio *top-down*, invece, gli algoritmi eseguiti dai singoli sensori utilizzano le informazioni provenienti da altri dispositivi. In questo caso l'output di ogni sensore dipende dalla presenza di ulteriori sorgenti di informazione nell'area circostante. Ognuno di essi implementa dunque una parte della logica di integrazione, realizzata ad un più basso livello di astrazione. Ad es. mediante un approccio top-down applicato all'analisi video, si può comprendere quando un cambiamento nella scena è riconducibile ad una ordinaria evoluzione della stessa (ad es. a causa di fattori metereologici) e aggiornare di conseguenza il background.

5. Processo di Information Fusion

Per garantire un adeguato livello di affidabilità, è fondamentale impiegare un approccio rigoroso, finalizzato a sfruttare i benefici offerti da tecnologie (smart-sensing), tecniche avanzate di sorveglianza (multi-modale e multi-mediale), e metodi di fusione delle informazioni. L'approccio al problema deve avere sia una base metodologica che una applicativa, al fine di trovare giusti compromessi in base alle dimensioni del flusso di informazioni da analizzare in tempo reale. Sono funzionali a tale scopo:

- La definizione di un'architettura per l'integrazione di tecnologie, tecniche e strumenti attualmente disponibili;
- Lo studio e l'applicazione di opportuni modelli e meccanismi di rilevamento.

Riguardo il primo punto, è necessario un paradigma finalizzato all'aumento delle capacità nella sorveglianza distribuita, a prescindere dal contesto applicativo. Esso risponde anche alla necessità di sfruttare la prospettiva dell'utente nella progettazione del sistema, e di migliorare l'intero processo decisionale.

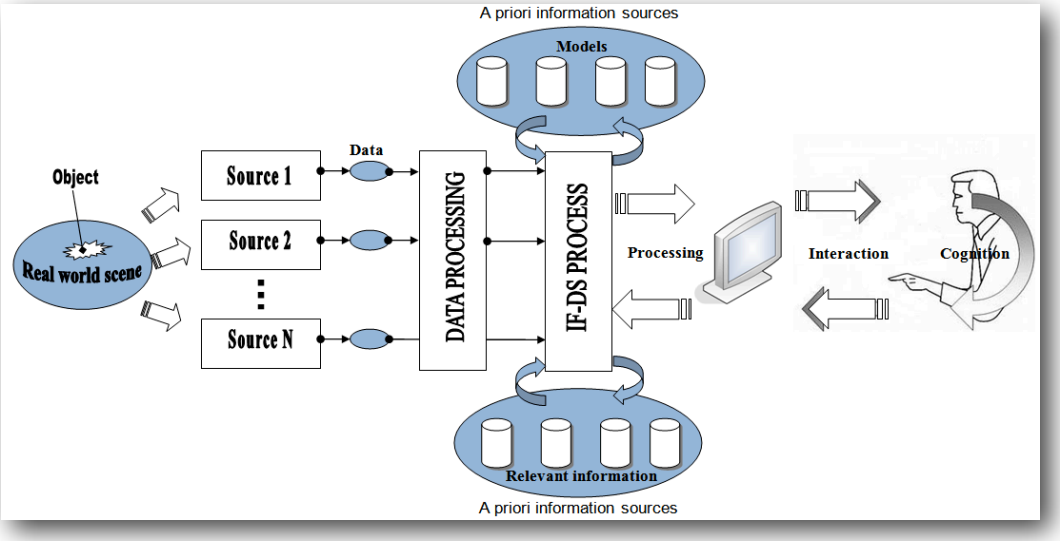

Figura 2
Un processo complessivo per il trattamento delle informazioni

La Figura 2 mostra il processo complessivo di trasformazione dell'informazione, a partire dai dati catturati dai sensori, fino ad arrivare ad indicazioni puntuali che un utente può usare per prendere decisioni. Esso si ispira ad uno dei più diffusi modelli di IF, denominato JDL (si veda il Riquadro 2) ed evidenzia tre importanti aspetti da tenere in considerazione nelle applicazioni di interesse:

a. La necessità di pre-elaborare i dati grezzi provenienti dai sensori installati sul campo. Ciò è legato al livello di eterogeneità dei sottosistemi di monitoraggio, che possono variare dai sensori di temperatura alle telecamere intelligenti, e al livello semantico dell'informazione da essi fornita. A seconda dei casi, la pre-elaborazione può essere presa in carico dai sottosistemi stessi, dai sistemi di integrazione e gestione di più tecnologie (che nel dominio della sicurezza fisica prendono il nome di sistemi PSIM, *Physical Security Information Management*), o anche immediatamente prima della fusione delle informazioni.

b. L'identificazione di differenti livelli di capacità (Figura 3). Infatti, la combinazione dei dati è finalizzata alla valutazione di: i) stati, attributi o identità delle singole entità monitorate; ii) le mutue relazioni tra le entità monitorate, anche rispetto all'ambiente circostante; iii) gli stati futuri e le proiezioni a partire dalla situazioni riconosciute, al fine di accertare il possibile accadimento di minacce e il relativo impatto.

c. La realizzazione di un costante processo di affinamento, il quale può essere guidato dall'utente e/o automatico. In particolare, l'utente può contribuire al processo di IF, supportando l'aggiornamento dinamico delle informazioni definite a priori, ovvero dei modelli e parametri della fusion. Questo affinamento è essenziale per consentire, ad esempio, l'uso di informazioni sempre in linea con le indicazioni provenienti dal campo.

Nella Figura 2, ogni sorgente (*Source*) può rappresentare un singolo sensore (a prescindere dal proprio livello di "intelligenza"), un più complesso sottosistema di monitoraggio (ad esempio multimodale), o ancora un operatore umano. Successivamente, in base al tipo di informazione fornita in uscita, può essere richiesta un'elaborazione preliminare (*Data Processing*), oppure l'output delle sorgenti può entrare direttamente come input del processo *IF-DS*. Il processo combinato di IF e DS considera come ulteriori input una base di conoscenza, rappresentata da modelli di rilevamento, e uno o più database, impiegati per registrare informazioni rilevanti come i feedback degli utenti (si veda il Riquadro 3). L'interazione con l'utente consente di conciliare il processo cognitivo automatico con quello umano, permettendo la comprensione di eventi e situazioni, il riconoscimento di fenomeni emergenti, e l'adozione di più precise contromisure.

Dal punto di vista applicativo, il paradigma presentato può essere implementato da un sistema integrato per la sorveglianza e la gestione di situazioni critiche, le cui capacità dovrebbero includere:

- l'interfacciamento con diverse piattaforme di monitoraggio;
- il rilevamento di eventi di interesse nell'ambiente monitorato;
- l'invio di segnalazioni agli operatori dei centri di controllo, in modo da supportare procedure di emergenza o anche l'attivazione di risposte automatiche.

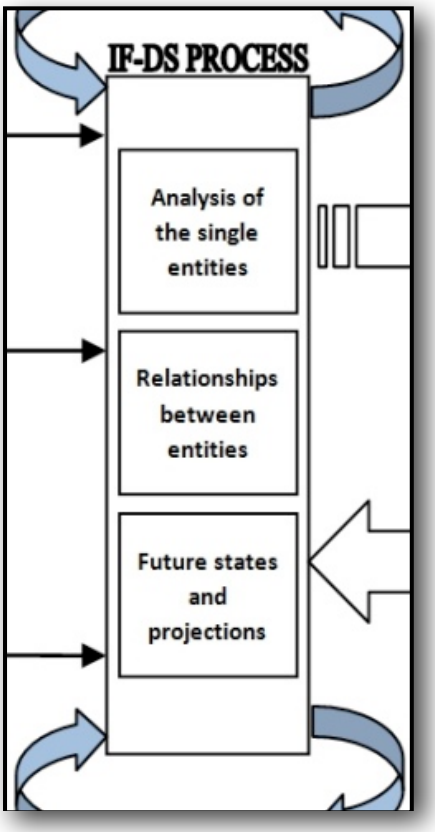

Figura 3
Dettaglio sul processo di IF e DS

Dal momento che il sistema è potenzialmente in grado di "inondare" gli operatori con una mole di informazioni ingestibile, la capacità di correlazione degli eventi è estremamente importante, per consentire agli operatori di concentrarsi sulle notifiche significative, ovvero indicative di potenziali minacce. Un framework integrato comprensivo di tutte le capacità appena descritte sarebbe dunque in grado di rilevare fenomeni di una certa complessità con elevata affidabilità ed efficienza. Il problema della correlazione di eventi è stato ampiamente studiato nella letterature scientifica ed un'ampia classe di soluzioni è stata definita. Tuttavia, i principali risultati hanno interessato applicazioni diverse dalla sorveglianza, come il rilevamento di intrusioni in reti informatiche o il riconoscimento di condizioni specifiche nelle cosiddette basi di dati attive. Nel contesto della sicurezza fisica, le capacità dei sistemi *legacy* sono molto limitate nell'analisi e nell'interpretazione dei dati in tempo reale. La scarsezza di applicazioni in questo settore è motivata dalla mancanza di approcci che siano al tempo stesso efficaci e facili da implementare: il riconoscimento di situazioni in evoluzione, basato su modelli di minaccia definiti a priori e facili da aggiornare anche da parte degli operatori, è un obiettivo tutt'altro che semplice da raggiungere. E tuttavia questa capacità è necessaria quando un elevato numero di entità monitorate è coinvolto in complesse relazioni spazio-temporali. Assumendo che gli scenari di minaccia siano decomponibili, durante l'analisi del rischio, in fasi elementari che si susseguono con un ordine abbastanza predicibile, è possibile

pensare ad una correlazione logica, spaziale e temporale degli eventi. Al contempo, occorre prevedere tecniche capaci di garantire la "resilienza" rispetto a possibili errori di modellazione o mancati rilevamenti (es. dovuti a guasti dei sensori). Da questo punto di vista, un approccio promettente consiste nell'impiego di tecniche di riconoscimento euristico, basate su analisi di similarità degli scenari rilevati con i modelli di minaccia definiti a priori [10].

6. Conclusioni e problemi aperti

In questo articolo si sono passate in rassegna alcune tematiche che ruotano intorno alla problema della integrazione, sintesi e gestione delle informazioni provenienti dai sistemi di sorveglianza. Come si è visto, si tratta di approcci che hanno avuto uno sviluppo relativamente recente e che sono per loro natura spiccatamente multidisciplinari, combinando ricerche in ambiti di *signal processing* a livello sensoriale, *multi-sensor information fusion* per sistemi distribuiti, intelligenza artificiale (classificazione, *pattern recognition* e sistemi esperti), modellistica predittiva di affidabilità e prestazioni del rilevamento, studi ergonomici e di *human-factors* sugli operatori.

L'importanza di un meccanismo di fusione delle informazioni efficace, efficiente e flessibile risiede nel fatto che esso risulterebbe applicabile ai più svariati sistemi di supervisione e monitoraggio diagnostico, usati in molteplici ambiti applicativi, dal momento che questi condividono lo scopo abbastanza generico di fondere dati sensoriali eterogenei per supportare le decisioni degli operatori.

La ricerca in quest'ambito ha ricevuto una notevole spinta dalla consapevolezza delle difficoltà di gestione dei sistemi SCADA (e similari) in applicazioni reali, per la mole ingestibile di allarmi ricevuti mediamente ogni giorno da ciascun operatore, che può arrivare a molte centinaia, di cui solo una piccola percentuale realmente significativi. Un altro fattore fortemente motivante è legato alla necessità di rendere "intelligenti" i sistemi di sorveglianza per la *homeland security*, al fine di controllare i comportamenti degli individui (anche nell'interazione con Internet) per prevenire eventuali attacchi (eventualmente di tipo *cyber*), senza al contempo violarne la *privacy*.

Un grosso contributo allo sviluppo di tecniche di sorveglianza avanzata è dovuto al trasferimento di conoscenze, modelli e tecnologie da altri ambiti, quali quello della difesa, degli *Intrusion Detection Systems* (IDS) per le reti di calcolatori oltre che del *data mining* e delle basi di dati attive. Ad oggi, purtroppo, nonostante le apparenti similitudini, nessuno di questi ambiti è riuscito a fornire qualcosa di facilmente, direttamente ed esaustivamente utilizzabile per gli scopi descritti in questo articolo.

Diversi sono i problemi ancora da risolvere nell'ambito della fusione delle informazioni applicata alla sorveglianza. Tra questi vale la pena citare la possibilità di specifica degli scenari tramite modelli ontologici (si veda ad es. il

riferimento [10]) facili da manutenere, che consentano una trasformazione automatica in opportuni modelli di rilevamento *real-time* la cui complessità è nascosta agli utenti finali. Un altro punto che merita approfondimenti è quello della gestione dell'incertezza/incompletezza nella rappresentazione delle situazioni di interesse e della potenziale inaffidabilità nel rilevamento sensoriale, che può essere basata su logiche *fuzzy* ed approcci euristici, supportati da taluni formalismi di modellazione. Infine, molto interessante è la problematica di gestione dell'apprendimento della struttura e/o dei parametri dei modelli di rilevamento, con un desiderabile aggiornamento dinamico in base ai feedback degli operatori (una sorta di *supervised learning*, non necessariamente basata su reti neurali o bayesiane).

Riquadro 1 – L'Information Fusion

La fusione delle informazioni**,** o ***information fusion***, è un'area di studio e di ricerca che ha ricadute estremamente significative in diversi domini applicativi, tra cui è possibile citare: difesa militare, previsioni metereologiche e finanziarie, diagnostica medica, allerta precoce di calamità naturali, etc. L'importanza che i processi e le tecniche di information fusion assumono nella progettazione e implementazione di una sempre maggiore varietà di sistemi è dovuta alla crescente mole di informazioni generata da sorgenti eterogenee (siano esse sensori, satelliti, social networks, o osservatori umani) e alla sempre più pressante necessità di trasformare tali informazioni in *conoscenza,* al fine di supportare i processi decisionali con sempre maggiore efficacia ed efficienza. L'information fusion richiede pertanto "*l'integrazione e la sinergia di informazioni circa il comportamento di uno specifico sistema, provenienti da sorgenti differenti, allo scopo di supportare decisioni e azioni relative al sistema stesso. L'information fusion include nel suo ambito teorie, tecniche e strumenti il cui obiettivo è sfruttare la conoscenza implicita nelle informazioni acquisite da più sorgenti e nelle relazioni tra di esse*" (fonte: www.isif.org). Diversi sono i convegni e le riviste che trattano tematiche attinenti, tra i quali il riferimento probabilmente più significativo è rappresentato dalla rivista *Information Fusion* di Elsevier, a cui si rimanda per ulteriori approfondimenti sullo stato dell'arte in ambito sia teorico che applicativo.

Riquadro 2 – Il modello JDL

Uno dei modelli più diffusi per classificare il processo di fusione nasce dal U.S. Joint Directors of Laboratories (JDL), da cui prende il nome [12]. Nato negli anni ’80 dal lavoro del Data Fusion Working Group, il modello JDL non descrive un flusso di attività, ma piuttosto individua categorie di funzioni che, a partire dai dati grezzi forniti dai sensori, consentono di pervenire ai diversi gradi di inferenza. Nella Figura 4, sono evidenziate le due principali categorie, cui appartengono funzioni divise in sottoclassi (o Livelli): la prima categoria è costituita dalle funzioni di *assessment* dei dati provenienti dalle sorgenti (Livelli da 0 a 3), la seconda da funzioni di *refinement*, alle quali è demandata la trasformazione delle informazioni in conoscenza (Livelli 4 e 5). Infine, il Livello 5 “apre” all’interazione tra il sistema di fusion e l’operatore umano, preposto a prendere le opportune decisioni in merito alle situazioni in essere (HCI: *Human-Computer Interaction*)

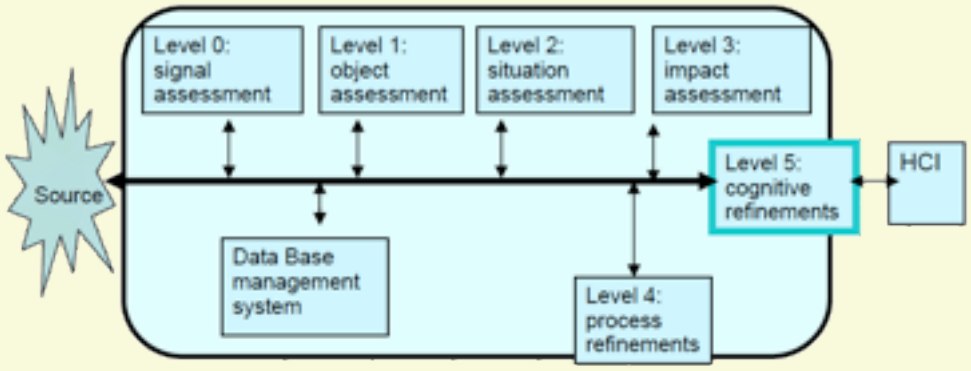

Figura 4
Il modello JDL

Riquadro 3 – Sorgenti e capacità nell'Information Fusion

Le informazioni relative al sistema reale che si sta osservando possono ricoprire un vasto orizzonte temporale. Come mostrato in Figura 5, in generale esse possono riferirsi al "presente" (i dati attuali provenienti, ad esempio, da sensori che stanno monitorando il sistema), al "passato" (attraverso la conoscenza pregressa, ad esempio serie storiche di eventi contenute in basi di dati), e al "futuro" (ad esempio, informazioni provenienti da simulazioni). Recentemente è stata riconosciuta l'importanza degli utenti nel contribuire efficacemente al processo di fusione delle informazioni.

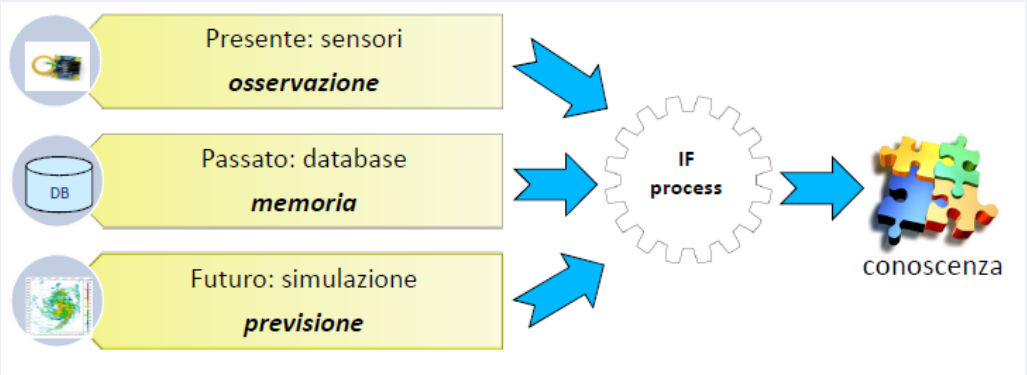

Figura 5
Schema di un sistema di Information Fusion

Il processo di fusione consiste in un progressivo meccanismo di inferenza che, a partire dai dati grezzi, permette di giungere: 1) alla determinazione dell'esistenza di entità di interesse; 2) alla loro localizzazione (posizione, velocità, etc.); 3) alla identificazione di tali entità; 4) alla definizione del loro comportamento; 5) alla comprensione della situazione in essere; 6) alla valutazione delle minacce che possono determinarsi dall'evoluzione futura (Figura 6).

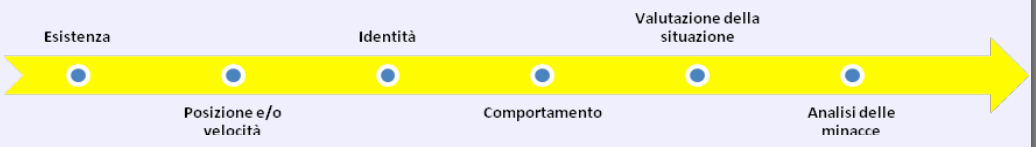

Figura 6
Livelli di inferenza

Riferimenti bibliografici

[1] Atrey, P. K., Hossain, M. A., El-Saddik, A., Kankanhalli, M. S., *Multimodal fusion for multimedia analysis: a survey*, Multimedia Syst., 16(6), 2010: pp. 345-379.

[2] Bocchetti, G., Flammini, F., Pappalardo, A., Pragliola, C., *Dependable integrated surveillance systems for the physical security of metro railways*, in: Proc. of Third ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC '09), Como (Italy), 2009: pp. 1-7.

[3] Bossé, E., Guitouni, A., Valin, P., *An Essay to Characterise Information Fusion System*, in: Proceedings of the International conference of Information Fusion. Florence, Italy, 2006.

[4] Castro, J.L., Delgado, M., Medina, J., Ruiz-Lozano, M.D., *Intelligent surveillance system with integration of heterogeneous information for intrusion detection*, in: Expert Systems with Applications, Vol. 38, No. 9, 2011: pp. 11182-11192.

[5] Cesarini, M., Ghezzi, C., Tanca, L., Schreiber, F., *Reti di microdispositivi intelligenti*, in: Mondo Digitale, n. 17, 2006: pp. 47-55.

[6] Cohen, J., Cohen, H. H., *Enhancing forensic human factors/ergonomics analyses using digital surveillance video*, in: Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting, 2007: pp. 1129-1132.

[7] Cranor, L. F., Garfinkel, S., *Security and Usability. Designing Secure Systems that People Can Use*, O'Reilly Media, 2005

[8] Cucchiara, R., *La visione artificiale per la videosorveglianza*, in: Mondo Digitale, n. 27, 2008: pp. 39-47.

[9] Flammini, F., Mazzocca, N., Pappalardo, A., Pragliola, C., Vittorini, V., *Augmenting Surveillance System Capabilities by Exploiting Event Correlation and Distributed Attack Detection*, in: Proc. of Intl. Workshop on Security and Cognitive Informatics for Homeland Defence (SeCIHD'11), co-located with ARES'11, Springer LNCS 6908, 2011: pp. 191-204.

[10] Flammini, F., Pappalardo, A., Pragliola, C., Vittorini, V., *A robust approach for on-line and off-line threat detection based on event tree similarity analysis*, in: Proc. of 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS'11), Klagenfurt, Austria, Aug. 30-Sept. 2, 2011: pp.414-419.

[11] Flammini, F., Pappalardo, A., Vittorini V., *Challenges and emerging paradigms for augmented surveillance*, in: Effective Surveillance for Homeland Security: Balancing Technology and Social Issues, CRC Press, 2013: pp. 169-198

[12] Hall, D.L., *Multisource information fusion for critical infrastructure situation awareness*, in: Critical Infrastructure Security: Assessment, Prevention, Detection, Response, WIT Press, 2012: pp. 267-277

[13] St. John, M., Risser, M. R., *Sustaining vigilance by activating a secondary task when inattention is detected*, in: Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting, 2009.

[14] Wickens, C., Dixon, S. *The benefits of imperfect diagnostic automation: a synthesis of the literature*, in: Theoretical Issues in Ergonomics Science, 8(3), 2007: pp. 201-212.

[15] Zhu, Z., Huang, T.S., *Multimodal Surveillance: Sensors, Algorithms and Systems*, Artech House Publisher, 2007.

Biografia

Francesco Flammini ha ottenuto la laurea con lode (2003) ed il dottorato di ricerca (2006) in Ingegneria Informatica presso l'Università di Napoli Federico II. Dal 2003 lavora in Ansaldo STS (Finmeccanica) come progettista e ricercatore, occupandosi prevalentemente di verifica dei sistemi di controllo, protezione delle infrastrutture, e Innovation Network. Ha tenuto come professore a contratto corsi di informatica ed ingegneria del software. E' autore di numerosi articoli scientifici pubblicati su riviste, libri e atti di congressi internazionali. Svolge attività editoriali per libri e riviste sul tema dei sistemi sicuri ed affidabili ed è nel comitato di programma di diversi convegni internazionali, tra cui SAFECOMP. E' un ACM Distinguished Speaker ed un IEEE Senior Member.
E-mail: francesco.flammini@ieee.org

Alfio Pappalardo ha conseguito nel 2013 il dottorato di ricerca in Ingegneria Informatica ed Automatica presso l'Università di Napoli Federico II, con borsa di studio finanziata da Ansaldo STS. Nel Dicembre 2008 ha conseguito la Laurea in Ingegneria Informatica presso la stessa università, e nell'Ottobre 2009 il Master in Homeland Security presso l'Università di Bologna Alma Mater Studiorum. Dal 2009 si occupa di attività di ricerca nei settori della protezione delle infrastrutture critiche e della security nei trasporti ferroviari e metropolitani, presso Ansaldo STS. Attualmente è coinvolto in progetti internazionali per soluzioni integrate di security in sistemi metropolitani.
E-mail: alfio.pappalardo@unina.it

Valeria Vittorini è professore associato presso l'Università di Napoli Federico II, dove è docente di Fondamenti di Informatica e di Programmazione. È autrice di numerose pubblicazioni scientifiche sull'utilizzo di tecniche formali di modellazione di sistemi reali in diversi ambiti applicativi. Afferisce al Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione, presso cui svolge la propria attività di ricerca, prevalentemente incentrata sullo studio di sistemi distribuiti e sullo sviluppo di metodologie e strumenti per l'analisi di sistemi critici.
E-mail: valeria.vittorini@unina.it