Gestire nuovi tipi di guasto ed anomalie emergenti su Infrastrutture Critiche altamente informatizzate

Claudio Balducelli*, Sandro Bologna*, Adam Maria Gadomski*, Luisa Lavalle*, Giordano Vicoli*
*ENEA – Ente per le Nuove Tecnologie, l'Energia e l'Ambiente
balducelli_c@casaccia.enea.it

SOMMARIO

In questo lavoro viene illustrato il progetto SAFEGUARD[1] che si propone di realizzare un sistema avente l'obbiettivo di proteggere le infrastrutture critiche mediante un insieme di agenti distribuiti dotati di algoritmi in grado di monitorare precocemente anomalie e malfunzionamenti dei componenti dell'infrastruttura ed attuare le più appropriate politiche di protezione. Il sistema SAFEGUARD trova applicazione principalmente per le infrastrutture altamente informatizzate quali le reti di telecomunicazioni e i sistemi SCADA[2] di supervisione e controllo delle reti elettriche; per questo tipo di reti complesse infatti le probabilità di guasto o attacco informatico risultano oggigiorno sempre crescenti, mentre sono difficilmente disponibili modelli formalizzati di comportamento e riconoscimento dei guasti stessi.

Nel lavoro verranno anche illustrati i risultati preliminari dei test effettuati in un ambiente di emulazione di un sistema SCADA di controllo e supervisione di una rete elettrica .

2. INTRODUZIONE

Il comportamento delle reti complesse, in presenza di guasti o anomalie risulta molto spesso differente da quello che ci si potrebbe ragionevolmente attendere e non sempre classificabile all'interno di regole o schemi logici precisi.

Man mano che la complessità ed il numero dei componenti di un sistema cresce, anche il numero delle possibili anomalie e la tipologia delle condizioni di guasto cresce in modo esponenziale. Al contrario, il numero dei comportamenti "normali" è generalmente più limitato. Ciò dipende dal fatto che una infrastruttura critica, come ad esempio una rete di distribuzione di energia, è progettata per fornire un servizio alla collettività che si esplica in modo generalmente sempre uguale; se variazioni vi sono esse dipendono da cause conosciute, generalmente ripetitive, quali la maggiore o minore richiesta di energia dovuta ad effetti periodici e stagionali .

In questo caso quindi, rispetto ai più tradizionali metodi diagnostici, sembra essere più efficiente un meccanismo in grado di "apprendere" il comportamento del sistema durante la sua normale operatività: il riconoscimento di "deviazioni" rispetto allo stato di funzionamento normale può in tal modo diagnosticare un eventuale condizione di guasto o di attacco del sistema (Balducelli et. al). Si potrebbero quindi riconoscere anche le situazioni "nuove" e non ancora sperimentate; il passo successivo sarebbe quello di analizzare e classificare in termini di grado di severità la situazione riconosciuta, al fine di stabilire le più appropriate politiche di protezione.

Anche nel settore più generale della gestione delle emergenze si sente oggigiorno la esigenza di avere una sensoristica intelligente in grado di riconoscere situazioni nuove e non aspettate. Specialmente negli ultimi anni infatti ci si è resi conto che il tipo di eventi che si devono affrontare sono totalmente "imprevedibili" ed inaspettati.

Basti pensare all'attacco nei confronti delle Torri gemelle: nessuna squadra preposta alla gestione delle emergenze (Esercito, Vigili del fuoco, Protezione Civile etc.) avrebbe mai potuto pensare di dover affrontare una simile situazione. La "creatività" e l'applicazione di nuove ed inusuali strategie (Kendra and Wachtendorf, 2002) si è resa assolutamente necessaria in quei frangenti.

Pensiamo poi agli attacchi chimici con sostanze tossiche spedite tramite il sistema postale che, a livello concettuale molto assomigliano alla spedizione di virus e "worms" tramite posta elettronica. Anche in questi casi ci si ritrova impreparati in quanto ciò che avviene è "inaspettato" e non vi sono procedure standardizzate per affrontare questo tipo di emergenza. Anche nel settore delle reti elettriche di trasmissione, negli ultimi anni, si sono verificate in modo inatteso e per cause non sempre evidenti, svariate condizioni di black-out sia in Europa che negli Stati Uniti.

[1] A multi-agent system to safeguard Large Complex Critical Infrastructures developed inside EU FP5 program
[2] Supervisory Control And Data Acquisition systems

Una caratteristica comune che sta emergendo in questi anni in tutte le Infrastrutture Critiche è la sempre maggiore diffusione al loro interno delle tecnologie Ict (Information and Communication Technologies) al fine di erogare più efficientemente servizi anche innovativi per andare incontro alle nuove aspettative ed esigenze dell'utenza; il *cyberspace* infatti sta assumendo il ruolo di *Global Information Infrastructure* ed è condiviso dalla maggior parte delle infrastrutture tecnologiche; a sua volta, l'esistenza del *cyberspace* è legata al corretto funzionamento di alcune di queste infrastrutture (rete elettrica, reti di telecomunicazione ecc.). Tutto questo viene mostrato sinteticamente in figura 1 nella quale risulta evidente come i legami tra esse si sviluppino prevalentemente sul piano del *cyberspace*.

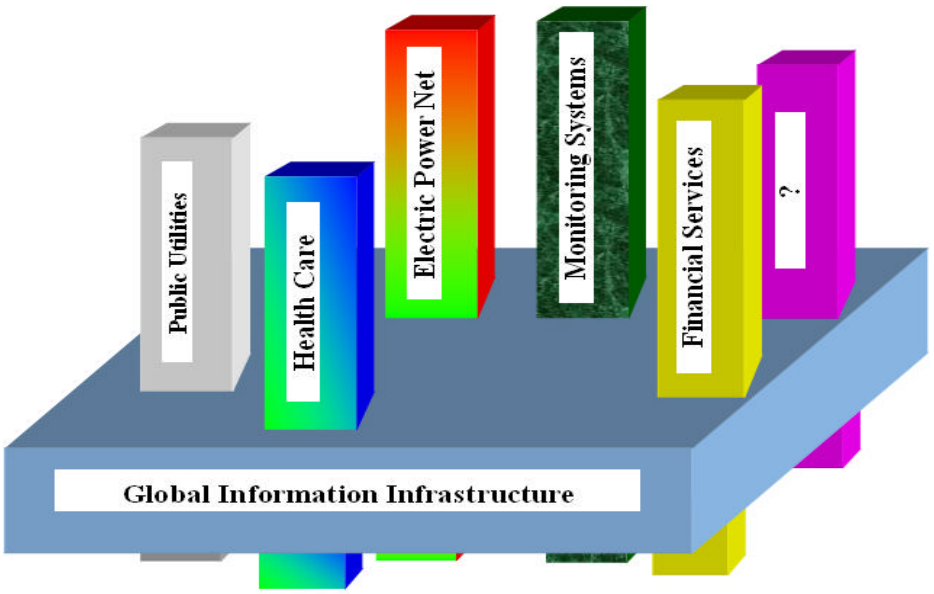

Figura 1 - Interdipendenza fra Infrastrutture Critiche

Il diffondersi delle tecnologie ICT introduce ed amplifica, però, l'interdipendenza fra le diverse infrastrutture critiche al punto che una malfunzionamento (accidentale o deliberato) di un'infrastruttura può ripercuotersi sulle altre provocando la diffusione del danno; questi problemi possono riflettersi anche su utenti remoti (sia dal punto di vista geografico che funzionale) rispetto al punto in cui esso si era verificato.

3. L'ARCHITETTURA SAFEGUARD

L'architettura del sistema SAFEGUARD è progettata con lo scopo di proteggere la rete informatica di controllo e supervisione nei confronti di attacchi, malfunzionamenti e stati anomali di funzionamento nei quali l'aspetto di imprevedibilità risulta molto marcato. SAFEGUARD è un sistema basato su agenti distribuiti che implementa un modello di *normalità di funzionamento* della rete e risponde in tempo reale all'insorgere di anomalie (deviazioni dallo stato di normalità).

In figura 1 è riportato lo schema della architettura ad agenti del sistema. Gli agenti SAFEGUARD di basso livello sono connessi con il "cyberlayer" della rete elettrica, che è un sistema SCADA di tipo distribuito. Eventuali problemi che vengono intercettati sul cyberlayer dagli agenti di *Hybrid Anomaly Detection (HDA)*, sono analizzati e diagnosticati meglio, a livello più alto dall'agente *Correlator*. Questo agente ha il compito di correlare queste anomalie con eventuali indicazioni acquisite dai sistemi più tradizionali di *Intrusion Detection* oltre che analizzare le dipendenze spaziali e temporali fra più eventi intercettati in punti diversi della infrastruttura.

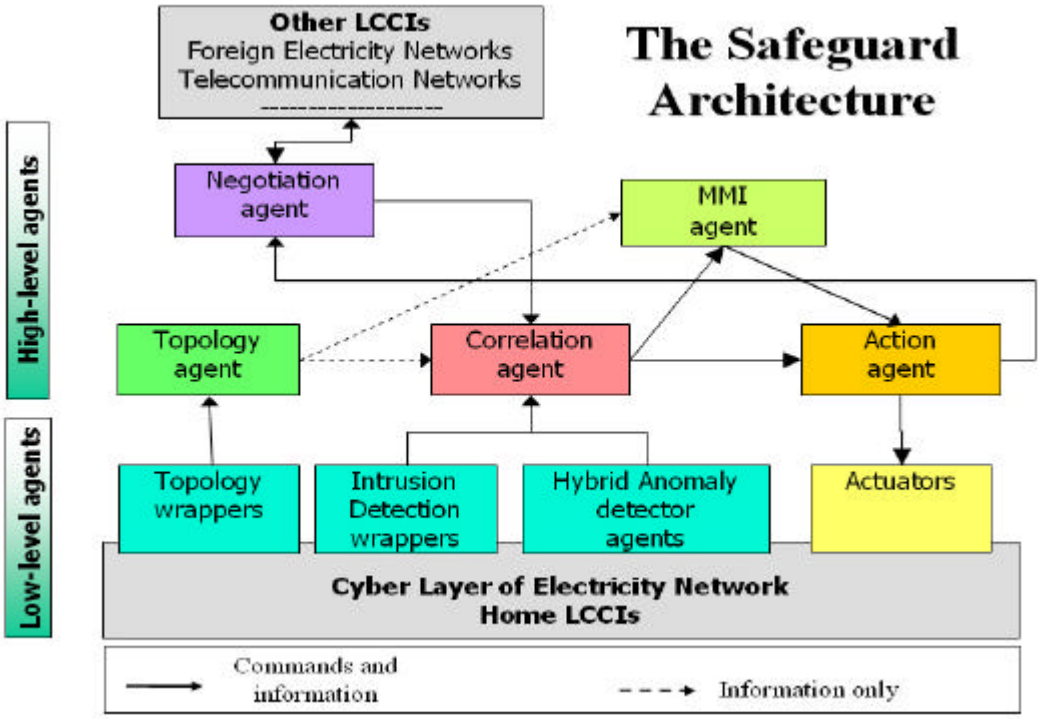

Fig. 1 – L'architettura ad agenti del sistema Safegurad

Il processo di correlazione viene supportato dall'agente *Topology*, che raccoglie su richiesta informazioni riguardanti la topologia della rete e dati di tipo statistico nei vari nodi.

Dopo una valutazione dello stato della rete, effettuata all'interno di opportune finestre temporali, l'agente *Action* produce una serie di risposte che includono sia dei comandi verso gli agenti di basso livello, sia delle azioni di salvaguardia della rete attraverso gli agenti *Actuators*.

Un agente *Negotiation* comunica con gli agenti istallati su altre Infrastrutture Critiche remote per richiedere i servizi e passare le informazioni circa gli allarmi che avvengono sulla infrastruttura locale e che possono avere delle interdipendenze verso l'esterno.

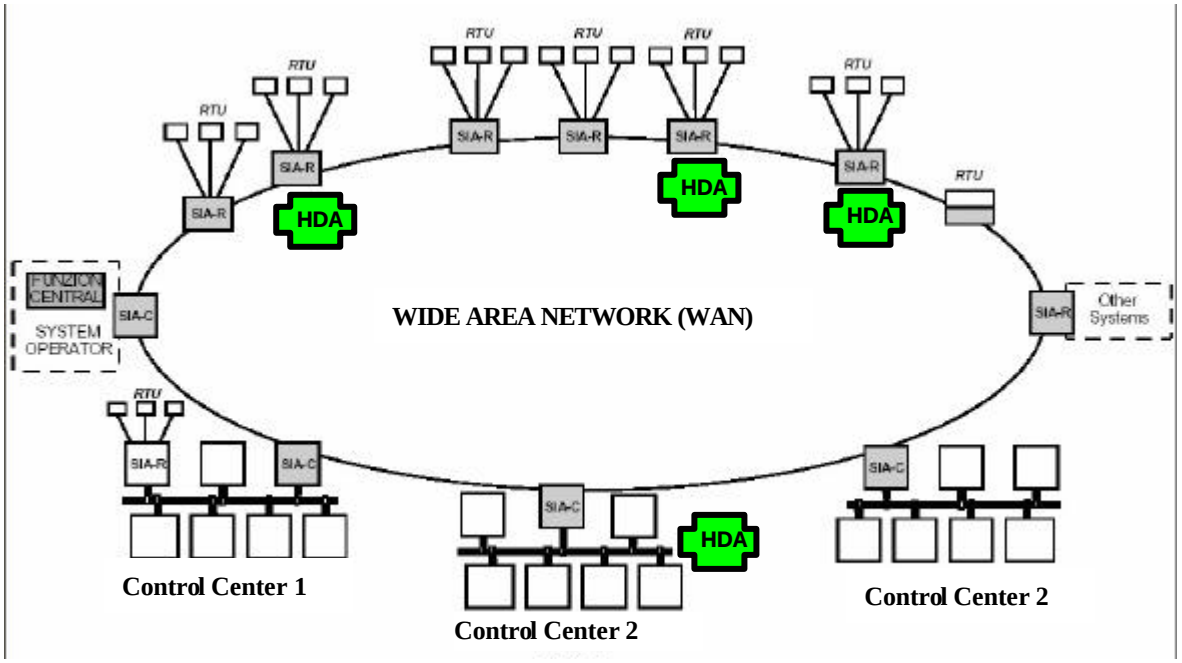

Fig 2 – Architettura del sistema SCADA della rete di trasmissione elettrica italiana

In Fig 2 è riportata la configurazione del sistema SCADA di dispacciamento e controllo dei flussi di potenza sulla rete elettrica italiana (M. Mocenico et al.). Dallo schema si vede che il sistema è disposto su una rete WAN (Wide Area Network) e comprende nodi di controllo centrale (Centri di Controllo), dotati al loro interno da più elaboratori connessi su rete locale e da grandi sale di controllo operativo, e da nodi di controllo periferico (SIA-R) che gestiscono una o più RTU (Remote Terminal Units). Le RTU sono i dispositivi digitali preposti allo scambio dati tra la rete elettrica ed i Centri di Controllo. Una o più RTU può essere in genere dedicata a gestire i dati provenienti da una singola sottostazione elettrica.

Nella figura è anche evidenziato come su alcuni nodi possono essere istallati agenti ibridi di anomaly detection (HDA – Hybrid Detection Agents)

3.1 Ruoli ed operatività degli agenti

Come evidenziato in fig 1, il sistema multi-agente SAFEGUARD è strutturato in due distinti livelli operativi: quello degli agenti di basso livello e quello degli agenti di alto livello.

Gli agenti di basso livello possiedono caratteristiche di velocità ed efficacia sia nella gestione delle informazioni in ingresso dal sistema SCADA che nella attuazione di azioni in uscita verso il sistema SCADA. Essendo il sistema di telecomunicazione e dispacciamento una infrastruttura distribuita su molti nodi, *più istanze* di questi agenti vengono prodotte e dedicate al monitoraggio ed alla analisi dei dati provenienti dai differenti nodi della rete. Gli agenti appartenenti a questo livello, che hanno una intelligenza piuttosto semplice e ottimizzata per poter processare in tempo reale le informazioni, vengono descritti qui di seguito.

Gli agenti *Wrappers* (WA), monitorano dati da differenti nodi di rete, e forniscono all'agente *Topology*, una interfaccia verso applicazioni tradizionali dedicate alla sicurezza quali i *firewalls*, i sistemi *antivirus*, ed altri sistemi di *intrusion detection*.

Gli agenti *Actuators* (ACA) hanno il compito di eseguire sul sistema SCADA I comandi e le politiche di difesa stabilite dall'agente *Action* con un eventuale consenso aquisito dall'operatore.

Gli agenti *Hybrid Detectors* (HDA) hanno il compito di monitorare la *normale attività* della infrastruttura, e di riportare all'operatore la visualizzazione di eventuali anomalie.

Varie tipologie di quest HDAs, come illustrato nel seguito, possono essere attivi su differenti componenti del sistema SCADA localizzati in punti differenti della rete. Ognuno di essi necessita normalmente di una fase di *addestramento* per riconoscere il modello comportamentale della rete.

Per gli agenti di alto livello sono sufficienti minori requisiti di velocità in quanto tutte le informazioni in ingresso sono già state filtrate dagli agenti a livello basso. L'agente Correlator, Action, MMI e Negotiation che agiscono su questo livello, sono responsabili di monitorare, analizzare e rappresentare informazioni critiche provenienti da vari agenti a basso livello, che non sono immediatamente analizzabili da questi agenti.

L'Agente *Correlator* (CA) ha il compito di eseguire una diagnosi precoce relativa al comportamento della rete, tramite la correlazione delle informazioni provenienti da differenti agenti di basso livello.

L'Agente *Action* (AA) riceve una informazione diagnostica dal Correlator ed attua, in più fasi successive al primo evento diagnosticato, una serie di azioni che rappresentano complessivamente una strategia di difesa avente lo scopo di riportare la rete verso uno stato di normalità. Per l'attuazione delle azioni più critiche e importanti dal punto di vista della gestione della rete può aspettare anche un consenso da parte dell'operatore.

L'Agente *Topology* (TA) include e gestisce la conoscenza relativa alla topologia del sistema SCADA e fornisce informazioni, su richiesta, al *Correlator*, all'*Action* ed al *MMI*.

L'agente *MMI* rappresenta l'interfaccia tra il sistema ad agenti e l'operatore della rete. Il suo ruolo principale è quello di fornire all'operatore tutte le informazioni rilevanti e filtrare quelle sovrabbondanti.

L'agente *Negotiation* si interfaccia con altre, simili o differenti, infrastrutture critiche con l'obbiettivo di minimizzare l'impatto di un malfunzionamento di una certa infrastruttura sulle altre.

3.2 Monitoraggio di anomalie

In SAFEGUARD vi sono tre distinti agenti ibridi di *Anonaly Detection* che hanno il compito di monitorare anomalie e guasti riscontrabili sul sistema SCADA:

- Un agente di riconoscimento delle principali sequenze di eventi di attivazione dei moduli software facenti parte del sistema SCADA. Esso viene indicato con la sigla ECHD (Event Courses Hybrid Detector) ed è basato basato su tecniche di Case Base Reasoning (CBR). Inizialmente esso apprende, fuori linea e per un tempo sufficientemente lungo, le relazioni temporali fra gli eventi appartenenti a queste sequenze (Terran & Brodley, 1999) ed associa un livello medio di anomalia ad ognuna di esse in funzione di quanto ogni sequenza intercettata di discosta statisticamente dalla media. Le sequenze di eventi che debbono essere riconosciute sono relative alla esecuzione delle differenti funzioni del sistema SCADA, e vengono inizialmente definite ed implementate nel Data Base usato dall'agente ECHD da un esperto del sistema SCADA stesso.
- Un agente di riconoscimento di deviazioni anomale dalla normalità della struttura e dei dati inclusi nei pacchetti che transitano attraverso le porte di comunicazione del sistema SCADA. Esso viene indicato con la sigla DMHD (Data Mining Hybrid Detector) in quanto utilizza vari algoritmi di Data Mining (Witten and Frank, 2000) per effettuare il riconoscimento.
- Un agente EDHD (Electricity Data Hybrid Detector) che associa un certo livello di anomalia, ai Data Sets di misure acquisite dal Sistema SCADA sulla rete elettrica. Le misure riguardano i valori di tensione, corrente, flussi di potenza attiva e reattiva. Queste misure debbono soddisfare determinate condizioni di invarianza. Se queste condizioni non sono verificate EDHD associa un grado di anomalia alle misure ritenute scorrette. Questo tipo di informazione viene utilizzato da uno strumento software di Stima dello Stato, generalmente disponibile come strumento di supporto operatore nei sistemi di supervisione e tele-controllo delle reti elettriche.

Un agente ibrido Safeguard possiede in genere due distinte funzionalità:

- la prima è dedicata a definire la *normalità* di funzionamento ed ad intercettare le deviazioni dal comportamento normale che potrebbero insorgere a fronte di guasti sconosciuti a priori;
- la seconda è dedicata a riconoscere la *firma* del comportamento e viene utilizzata dall'agente Correlatore al fine di generare allarmi relativi a malfunzionamenti della infrastruttura.

La prima è dedicate al riconoscimento di situazioni nuove e non ancora classificate, mentre la seconda ha natura di tipo diagnostico e tenta di rispondere in tempo reale ad una causa di guasto.

Nel prossimo paragrafo viene descritto con più dettagli l'agente ECHD per il riconoscimento di sequenze di eventi sui sistemi SCADA.

4. L'AGENTE IBRIDO ECHD

L'agente ECHD (Event Course Hybrid Detector) classifica e riconosce le sequenze di eventi all'interno del dei nodi che compongono il sistema di supervisione e controllo della rete elettrica, ossia i nodi del sistema SCADA, che nel nostro caso sono i Centri di Controllo e le RTU. Per definire ed implementare il modello delle sequenze di eventi più comuni di un sistema SCADA è stata usata una Base di Casi costruita seguendo la metodologia CBR (Case Base Reasoning).

L'approccio CBR parte dall'ipotesi che mediante una procedura di tipo algoritmico si può riconoscere solo un set limitato di comportamenti mentre nel mondo reale si ha la necessità di gestire anche situazioni e comportamenti non previsti più o meno dissimili da quelli conosciuti.

I primi sistemi CBR erano sistemi passivi: richiedevano cioè utenti che li attivassero manualmente e fornissero informazioni circa la natura dei problemi emergenti. Durante gli anni più recenti i tradizionali sistemi CBR passivi sono stati migliorati mediante l'introduzione di proprietà attive (Li and Yang, 2001). La nuova generazione di sistemi CBR ha capacità di risposta in tempo reale e flessibilità nella gestione della conoscenza così come un alto grado di autonomia nelle risposte rispetto ad un sistema CBR passivo.

La tecnica CBR è normalmente utilizzata per modellare, attraverso attributi eterogenei, un insieme di oggetti, chiamati "Casi", che rappresentano il dominio di intervento sul quale l'agente opera. La caratteristica

principale è quella della capacità dell'agente di riconoscere una maggiore o minore *somiglianza* di un Caso rispetto ad un altro sulla base di una *metrica* applicata sugli attributi che lo caratterizzano.

Nella applicazione Safeguard i Casi sono le sequenze di eventi che in un sistema SCADA rappresentano in sostanza i processi informativi associati alle varie funzioni espletate dal sistema stesso.

In figura 3 è riportata la sequenza di eventi relativa alla acquisizione e al processamento da parte del sistema SCADA di un tele-segnale. I tele-segnali rappresentano degli eventi asincroni che vengono acquisiti dalla rete elettrica: la chiusura o l'apertura di un interruttore, per esempio, generano un tele-segnale che deve essere acquisito dal sistema SCADA altrimenti l'operatore non riuscirebbe ad avere più una visione corretta ed aggiornata sullo stato della rete.

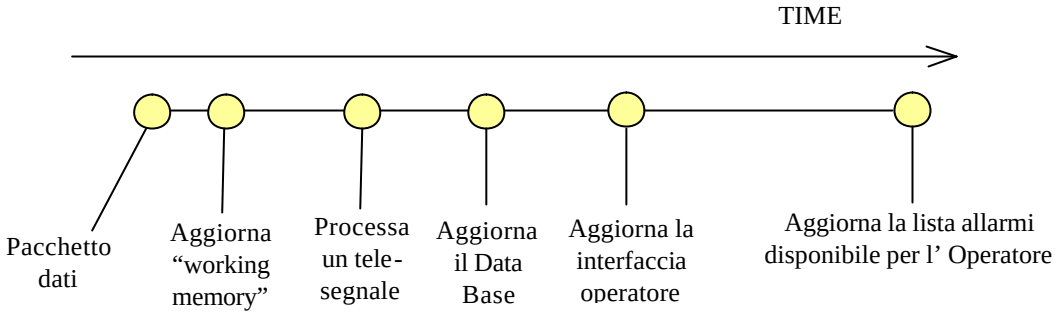

Fig. 3 – La sequenza di eventi relativa alla acquisizione e processamento di un tele-segnale

4.1 Struttura di un generico Caso

Volendo rappresentare questa sequenza come un Caso occorrerà definirne come attributi sia le tipologie di eventi in essa contenuti che i valori di ritardo temporale aspettato tra un evento e l'altro.

In modo più generale le sequenze di eventi verranno formalizzate come evidenziato nella Tabella 1. In questo caso la sequenza di eventi Es-E6 è contenuta all'interno di un intervallo temporale T; Es rappresenta il primo evento, o evento iniziatore della sequenza che avviene al tempo t=0; agli eventi successivi corrispondono delle etichette temporali (t1,..,t6) che rappresentano il ritardo medio dell'evento rispetto all'evento iniziatore, e delle deviazioni temporali (d1,..d6) che rappresentano le massime deviazioni passibili dell'evento dalla media.

Id del Caso: X	E_S	E_1	E_2	E_3	E_4	E_5	E_6	Tempo totale T
Etichetta temporale	0	t_1,d_{t1}	t_2,d_{t2}	t_3,d_{t3}	t_4,d_{t4}	t_5,d_{t5}	t_6,d_{t6}	

Tab. 1 – representation of a generic course of events Case

4.2 Definizione di una metrica per il confronto dei Casi

Durante la sua normale attività l'agente ECHD confronta tutte le sequenze di eventi memorizzate nella Base dei Casi con la sequenza che viene intercettata dal Sistema SCADA

Nella tabella 2 è riportata, a titolo di esempio, la sequenza intercettata in tempo reale. In essa si riconosce, oltre ad ulteriori eventi indicati con "*", anche la sequenza di tabella 1.

Sequenza corrente	E_S	*	*	E_1	E_2	*	*	*	E_3	*	E_4	E_5	*	E_6	*	*
Etichetta temporale	0	*	*	tc_1	tc_2	*	*	*	tc_3	*	tc_4	tc_5	*	Tc_6	*	*

Tab. 2 – La sequenza di eventi corrente acquisita in linea dal sistema SCADA

Essa viene perciò riconosciuta dall'agente ECHD, il quale comunque associa ad essa un livello di anomalia che corrisponde alla distanza D della sequenza stessa con quella memorizzata nella Base di Casi tramite la

seguente metrica:

$$D = \frac{\sum_{i=1}^{6} \frac{|t_i - tc_i|}{d_{ti}}}{6} \quad (1)$$

$$D = \frac{\sum_{i=1}^{n} \frac{|t_i - tc_i|}{d_{ti}}}{n} \quad (2)$$

Nel caso di sequenze è composte da un numero n arbitrario di eventi la (1) può essere scritta come in (2).

La distanza D viene normalizzata ad uno ed assume i valori compresi tra 0 ed 1 e rappresenta il livello di anomalia dell'intera sequenza.

Si può poi considerare esclusivamente il livello di anomalia di un singolo evento i della sequenza che è dato da:

$$D_i = \frac{t_i - tc_i}{d_{ti}} \quad (3)$$

L'analisi contemporanea dei livelli di anomalia sull'intera sequenza e su un singolo evento fornisce le seguenti importante informazioni aggiuntive:

- un alto livello di anomalia associato all'intera sequenza in assenza di livelli particolarmente alti di anomalia associati ad un singolo evento, è indicatore di un "malfunzionamento diffuso" associabile a rallentamenti o disturbi di tipo generale.
- un alto livello di anomalia associato al singolo evento, anche in presenza di un livello piuttosto basso di anomalia associato all'intera sequenza, è indicatore di un malfunzionamento associato ad un evento particolare che può essere causato da un "disturbo o guasto specifico" di un singolo componente del sistema SCADA.

4.3 Training dell'agente ECHD

Una fase di training è necessaria per inizializzare o aggiornare le etichette temporali delle sequenze di eventi. Inizialmente le sequenze di eventi, a cui vengono associate etichette temporali di comodo, vengono definite dagli esperti del sistema SCADA mediante l'utilizzo di un "Case Base editor".

Durante la fase di training l'agente ECHD non esegue alcun tipo di diagnostica; monitora esclusivamente le sequenze di eventi per un tempo sufficiente ad ottenere insiemi di dati atti ad eseguire su di essi le funzioni statistiche necessarie per ottenere appropriate medie e varianze delle etichette temporali.

5. CORRELAZIONE SPAZIO-TEMPORALE

In una infrastruttura complessa, come un sistema SCADA distribuito, composto da una rete di nodi di elaborazione fra loro interconnessi, le funzioni dell'intero sistema non si esplicano solo all'interno di un singolo nodo ed di un singolo intervallo temporale. Le funzioni sono espletate sia a livello spaziale che temporale.

Per questo motivo così come una anomalia evidenziata in singolo nodo da un agente ECHD descritto nel precedente paragrafo, non può da sola essere indicativa di un malfunzionamento generale della intera infrastruttura, anche l'assenza di anomalie in un certo nodo può non essere da sola essere indicativa del fatto che nessun malfunzionamento sia "in progress".

Soltanto attraverso la "correlazione" di ciò che sta avvenendo in un certo nodo con quanto sta avvenendo

in un altro si possono intercettare fenomeni che per loro natura non sono localizzabili in un solo punto della rete.

Consideriamo ad esempio il seguente processo. L'operatore del sistema SCADA invia un tele-comando dal Centro di Controllo avente l'obbiettivo di aprire l'interruttore di una connessione di linea sulla rete elettrica. Il telecomando viene inizialmente confezionato dal software del sistema all'interno di un pacchetto e attraverso la rete di telecomunicazione parte dal Centro di Controllo e transita in una Unità Remota (RTU) dislocata sulla porzione di territorio ove si trova la connessione da aprire. La RTU processa il tele-comando e quindi lo invia sulla rete elettrica ed un servo meccanismo che consente l'apertura dell'interruttore.

Per verificare che non vi siano malfunzionamenti sulla catena di attuazione del tele-comando occorre verificare che in un intervallo temporale seguente il momento in cui il telecomando viene processato dal centro di controllo esso transiti e sia processato sulla RTU; altrimenti significherebbe che il telecomando è andato perduto e vi è quindi un malfunzionamento nella catena di trasmissione dei pacchetti. In altre parole occorre correlare, a livello spazio-temporale fenomeni che avvengono sui diversi nodi della rete.

Questa funzione diagnostica viene espletata dall'agente Correlatore tenendo conto delle segnalazioni inviate dagli agenti ECHD, che come evidenziato in fig. 1, sono istallati su vari nodi della rete elettrica.

6. GESTIONE DEI FALSI ALLARMI

Una delle più importanti difficoltà dei sistemi basati sulla "detection" di anomalie, soprattutto su infrastrutture altamente informatizzate, è la gestione dei falsi allarmi. Infatti alcune deviazioni dalla condizione di normale funzionamento possono a volte insorgere non per una reale condizione di guasto o per un attacco realmente effettuato ma bensì per congestioni dovute ad un temporaneo sovraccarico della infrastruttura o a temporanee modifiche dello stato rete che non erano previste a priori e che quindi non rientrano tra le condizioni di normale funzionamento.

Nel sistema Safeguard, mentre gli agenti di basso livello hanno il compito di intercettare gli stati anomali come sintomi precoci di possibili malfunzionamenti più generali, gli agenti di alto livello, e specialmente il Correlatore e l'Action agent, hanno il compito di analizzare in un arco temporale più ampio l'insieme di anomalie intercettate, riducendo in tal modo la possibilità di reazioni inadeguate.

Questo comportamento è simile a quello di una organizzazione complessa nella quale un manager di alto livello collabora con responsabili di più basso livello.

Consideriamo, a titolo di esempio, ciò che accade in una organizzazione complessa che opera nel campo del trasposto merci. Un responsabile ad alto livello ha il compito di controllare la efficacia del trasporto delle merci attraverso un sistema di smistamento e consegna delle merci stesse composto da molti nodi o piattaforme di transito ("hubs"). Molti responsabili di più basso livello devono invece controllare il processo di smistamento delle merci attraverso le singole piattaforme di cui essi sono i gestori locali. Quando un responsabile di una singola piattaforma si accorge di un anomalia, ad esempio una congestione o un rallentamento nel carico e scarico delle merci all'interno della piattaforma sotto il suo controllo, egli comunica questa anomalia al responsabile di più alto livello. Quest'ultimo, pur prendendo in considerazione la anomalia, prima di prendere decisioni drastiche quali chiudere la piattaforma o dirottare le merci su altre piattaforme, fa le seguenti considerazioni:

- Analizza se la congestione possa essere considerata come una anomalia locale magari risolvibile nel prossimo futuro, anche tenendo conto del livello di congestione presente sugli piattaforme limitrofe.
- Guadagna un po' di tempo, con l'obbiettivo di analizzare la situazione durante un ulteriore intervallo temporale nel quale potrà verificare se la situazione anomala in crescita o in decrescita e se sia più o meno persistente

Solo tramite questi dati addizionali, acquisiti in un arco di tempo successivo al primo insorgere dell'anomalia, egli può decidere se sia necessaria o meno attuare una politica di emergenza su scala più vasta o se la situazione debba solo essere tenuta sotto controllo in quanto vi è la forte probabilità di poter essere risolta in breve tempo a livello locale.

Nel sistema SAFEGUARD tali politiche di emergenza vengono attuate dall'Action agent sulla base degli allarmi forniti dal Correlator agent. In particolare l'Action agent utilizza la metodologia dei Workflows per formalizzare ed eseguire tali politiche. In fig. 4 è visualizzato a titolo di esempio il Workflow relativo alla gestione di una situazione di crisi causata dalla disconnessione indesiderata di una linea elettrica e dal fallimento nella esecuzione del telecomando inviato per richiudere il relativo interruttore. Nella sequenza si può notare come politiche di recovery sempre più rafforzative vengono eseguite, con il consenso dell'operatore,

per tentare di attuare il riconnessione della linea. Se ciò non si riesce con le azioni previste si prevede di informare e chiedere aiuto, tramite il Negotiation agent agli operatori delle reti di trasmissione elettrica confinanti.

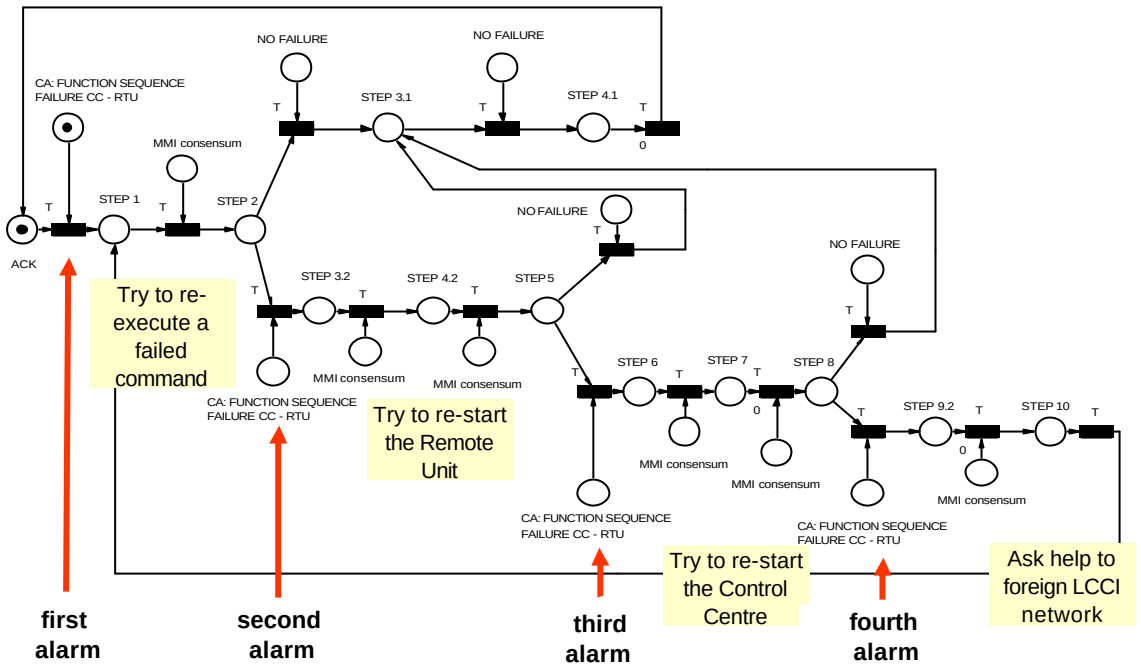

Fig 4 – Workflow di gestione di malfunzionamenti nell'invio di tele-comandi

7. L'EMULATORE SCADA

L'ambiente di test del Sistema Safeguard è composto da:

- Un simulatore di "loadflow" di rete elettrica (e-Agora)
- Un emulatore di un Sistema SCADA (ES)

Il simulatore della rete elettrica (eAgora) è stato realizzato dall'AIA[3]. L'emulatore del Sistema SCADA è stato completamente realizzato in ENEA ed ha il compito di fornire i dati di loadflow al Sistema Safeguard in due modalità diverse:

- in maniera OFF-LINE, previa memorizzazione su database per la fase di addestramento dei vari agenti;
- in maniera ON-LINE durante la fase operativa del Sistema Safeguard stesso.

L'ES interagisce con il simulatore della rete elettrica e-Agora. In particolare acquisisce ciclicamente le TeleMisure, invia i TeleComandi di apertura/chiusura interruttori, riceve i TeleSegnali.

Le TeleMisure, sono "valori continui" di intensità di corrente, angoli, etc. calcolati dal simulatore e-Agorà. Tale calcolo viene effettuato su richiesta dell'ES e sulla base dell'invio del valore del carico. Il valore del carico viene calcolato sulla base di alcune tabelle appositamente preparate, varia ogni ora ed ha ciclicità annuale.

I Telesegnali sono "valori booleani" o discreti che vengo inviati in modo asincrono al sistema SCADA quando cambia stato di un certo apparato (p.e. un interruttore può trovarsi nello stato "chiuso" o "aperto"). Poiché il simulatore e-Agorà non genera automaticamente i Telesegnali, lo stesso ES simula questa caratteristica.

Il TeleComando consiste in una richiesta che viene effettuata dall'operatore attraverso un interfaccia uomo-macchina. L'operatore seleziona una sottostazione, un bus ed un interruttore. L'ES visualizza lo stato dell'interruttore selezionato e l'operatore decide se chiuderlo a aprirlo a seconda dello stato attuale. Tale richiesta dell'operatore si traduce in una richiesta dell'ES verso il simulatore della rete elettrica che esegue il

[3] Applicaciones de Informatica Avanzata – Barcelona

comando vero e proprio.

7.1 Architettura dell'Emulatore SCADA

Sono stati connessi su rete locale un insieme di elaboratori che emulano il comportamento della architettura SCADA di illustrata precedentemente tramite la fig. 2. L'ES è sostanzialmente una applicazione distribuita costituita da diversi componenti visualizzati in fig. 5. Ciascun componente può risiedere in un nodo diverso di una rete. Abbiamo un Centro di Controllo (CC) costituito da una GUI (interfaccia uomo-macchina) e da un DAC (modulo acquisizione dati). Il CC acquisisce i dati dalle SIA-R. Esse sono dei concentratori dei dati che provengono dalle RTU. Nell'ES le SIA-R possono essere un numero qualsiasi e ciascuna controlla un certo numero di sottostazioni dell'intera rete elettrica controllata dal simulatore eAgora.

Le SIA-R acquisiscono i dati attraverso il componente AD (Analogic-Digital Converter). Nel caso dell'ES il componente AD ha il compito di interagire con il simulatore e-Agora; invia i valori del carico per ogni ora simulata, visualizza su un grafico l'andamento del carico ora per ora, richiede il calcolo del loadflow, rende disponibili i risultati alle SIA-R, invia le richieste di TeleComando che provengono dal CC, genera i TeleSegnali sulla base dei risultati che provengono dal simulatore e-Agorà che invia al CC.

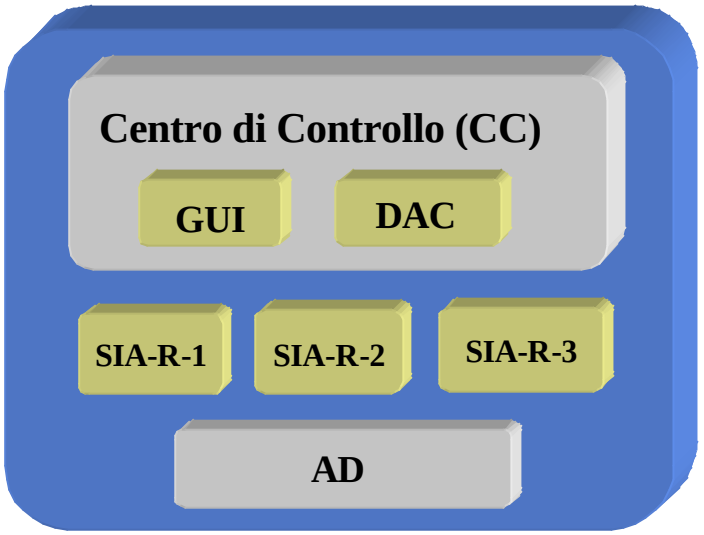

Fig.5 Architettura dell'Emulatore SCADA.

Ciascun componente dell'ES è stato realizzato utilizzando il linguaggio Java e comunica con gli altri componenti mediante messaggi. Per lo scambio dei messaggi è stata utilizzata la tecnologia JMS (Java Message Service) pertanto è stato necessario l'utilizzo di un broker di messaggi che è parte integrante dell'intero sistema. Nella figura 6 possiamo vedere come si inserisce il broker nell'intera architettura. Sostanzialmente ciascun componente dell'ES comunica con l'altro tramite questo server di messaggi che ha il compito di instradare i le informazioni al giusto destinatario attraverso un sistema di code.

Vi sono due modalità di invio dei messaggi. La prima, chiamata Point-To-Point, prevede l'utilizzo di una coda dove possono essere inviati dei messaggi. Vi possono essere più riceventi in ascolto su questa coda ma solo uno di essi riceverà il messaggio.

L'altra modalità, Publish-Subscribe, prevede l'invio dei messaggi verso una destinazione chiamata Topic. Tutti i riceventi che abbiano effettuato una sottoscrizione al topic riceveranno lo stesso messaggio. Tale sottoscrizione può anche essere duratura. Ciò significa che se un ricevente si disconnette momentaneamente quando si riconnetterà riceverà tutti i messaggi che sono arrivati durante la sua disconnessione. L'utilizzo di tale tecnologie comporta numerosi vantaggi. Il più importante riguarda il completo disaccoppiamento dei vari componenti. Se un componente invia un messaggio ad un altro componente quest'ultimo non necessariamente deve essere attivo o pronto a ricevere il messaggio. Il livello di astrazione con cui si lavora è alto per cui la produttività è anch'essa elevata.

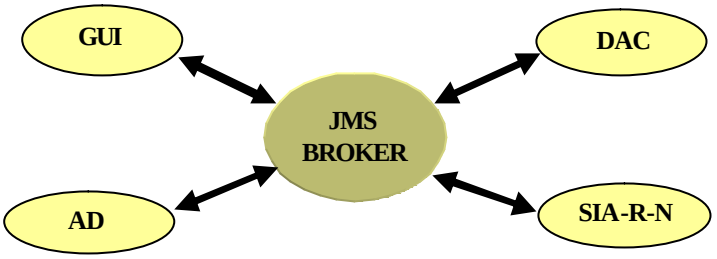

Fig.6 - Integrazione dell'ES con il broker di messaggi.

7.2 Funzione di "instrumentation"

Per monitorare lo stato del Sistema SCADA Safeguard deve controllare almeno in parte sia gli eventi che avvengono in esso sia il set dati provenienti dalla rete elettrica. Si parla quindi di *instrumentation* dell'ES in quanto l'ES stesso rende disponibili queste informazioni verso l'esterno. A tal fine è stata utilizzata la stessa tecnologia JMS impiegata per la realizzazione dell'ES. In particolare sia gli eventi che i dati vengono inviati ad un Topic in modalità Publish/Subscriber. In questo modo essi saranno disponibili per più Agenti Safeguard contemporaneamente.

Per quanto riguarda i dati elettrici essi non sono trattati come gli eventi. Infatti vengono resi disponibili al Sistema Safeguard ogni qual volta viene acquisita una TeleMisura. In particolare si hanno due modalità: la prima è quella in cui vengono resi disponibili i dati elettrici relativi all'intera rete elettrica; la seconda è quella in cui ciascuna SIA-R rende disponibili i dati relativi alle rispettive sottostazioni. I dati relativi alla rete elettrica sono in formato XML. E' possibile vedere come funziona il meccanismo di "instrumentation" dei dati elettrici nella figura 7 dove è illustrato il diagramma di sequenza di una richiesta di TeleMisura. Si può notare che una volta che la TeleMisura è stata calcolata essa viene inviata alla SIA-R e ai componenti DAC e GUI. Sia la SIA-R che il DAC inviano i dati elettrici al Topic. La SIA-R invia i dati relativi alle sue sottostazioni mentre il DAC i dati complessivi dell'intera rete elettrica e quindi di tutte le sottostazioni.

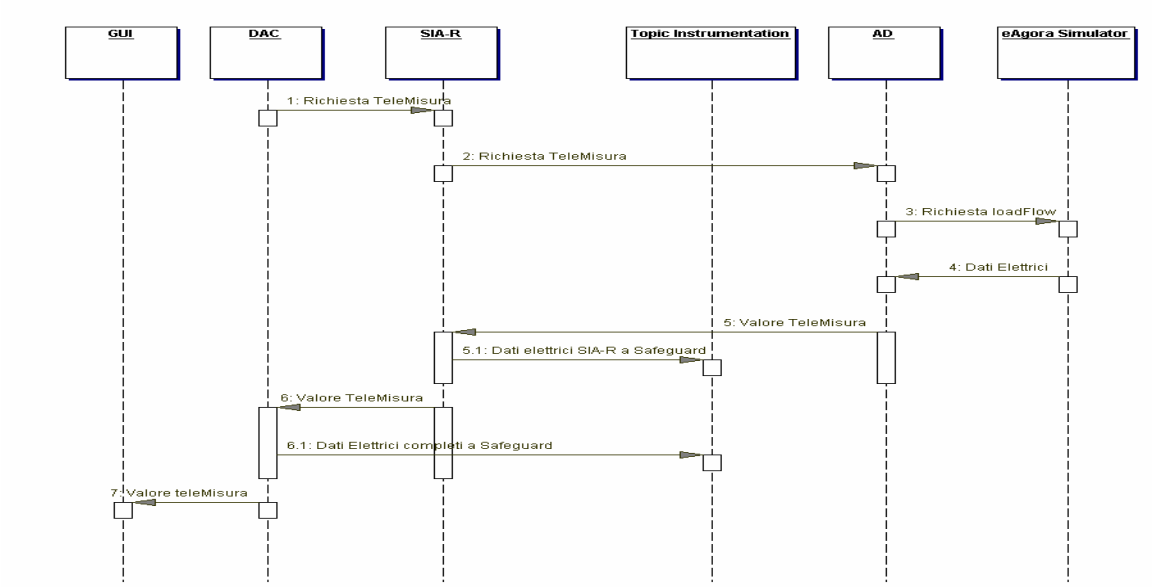

Fig. 7 - Diagramma di sequenza per l'instrumentation dei dati elettrici.

7.3 Utilizzo dell'emulatore SCADA per il test del sistema Safeguard

Per poter testare il Sistema Safeguard, nell'ES si è introdotta la capacità di simulare qualche malfunzionamento. Infatti l'ES è stato progettato in maniera tale da poter essere danneggiato sia dall'esterno con quelli che vengono chiamati *attacchi o comandi maliziosi contro il Sistema SCADA*, sia dall'interno con opportuni settaggi che l'utente può impostare in maniera molto semplice tramite menu. Attualmente l'ES può accettare

dall'esterno i seguenti attacchi e/o comandi:

- Invio di un falso TeleComando con dati non corretti (che non andrà a buon fine)
- Invio di un falso TeleComando con dati corretti (che andrà a buon fine)
- Invio di un corretto TeleComando al Centro di Controllo
- Invio di messaggio per la disabilitazione di una SIA-R al ricevimento di TeleComandi
- Invio di un messaggio per la re-inizializzazione di una SIA-R

Tramite menu l'utente può inoltre effettuare le seguenti impostazioni:

- Aumentare e/o diminuire i ritardi temporali degli eventi relativi alle TeleMisure
- Aumentare e/o diminuire i ritardi temporali degli eventi relativi ai TeleComandi
- Disabilitare l'invio dei TeleSegnali al Centro di Controllo
- Simulare problemi del Sistema Elettrico (ad esempio blackout) inviando soluzioni nulle del Simulatore e-Agora.

7.4 Scenari di test

Le funzionalità del sistema Safeguard sono state testate eseguendo sull'emulatore SCADA un insieme di scenari di test composti da attacchi o malfunzionamenti definiti e simulati mediante un "attack tool" che invia verso l'emulatore i comandi descritti nel precedente paragrafo.
Le differenti classi di scenari utilizzati per il test sono le seguenti:

- Invio di falsi tele-comandi e perdita di tele-comandi, con la conseguenza di invalidare le informazioni disponibili dall'operatore sul pannello allarmi.
- Corruzione di dati elettrici e/o modifica anomala dello stato dei componenti, con la conseguenza di invalidare le informazioni disponibili sulla Consolle operatore.
- Generazione di attacchi legati a Worms di natura sconosciuta, con la conseguenza di rallentare le sequenze operative nel sistema SCADA, e di produrre congestioni delle reti WAN e LAN.
- Generazione di perturbazioni e/o corruzione nei pacchetti di dati che transitano attraverso le porte di comunicazione utilizzate dall'emulatore SCADA.

Questi scenari sono composti da varie sequenze di azioni maliziose che vengono eseguite in maniera automatica dall'attack tool. Gli scenari sono stati suddivisi in più varianti distinte aventi insiemi diversificati di azioni, sfruttando la capacità dell'attack tool di generare patterns diversificati di azioni con differenti vincoli temporali fra di esse.
I risultati dei tests hanno evidenziato in modo soddisfacente le capacità degli agenti Safeguard di intercettare precocemente le situazioni anomale e di inviare gli allarmi adeguati sulla Consolle operativa.

8. BIBLIOGRAFIA

[1] C. Balducelli, L. Lavalle, G. Vicoli,Novelty Detection and Management to Safeguard Information Intensive Critical Infrastructure, *10th International Conference on Emergency Management*, Melbourne, Yarra Ranger, May 18-21 2004

[2] Kendra J., Wachtendorf T.(2002),Creativity in Emergency Responce after the World Trade Center Attack, *9th Annual Conference of The International Emergency Management Society*, Waterloo, Canada, May 14-17, 2002.

[3] Terran L., Brodley C. E.(1999), Temporal sequence learning and data reduction for anomaly detection, *ACM Transactions on Information and System Security (TISSEC)*, v.2 n.3, p.295-331, Aug. 1999

[4] Witten H., Frank E.(2000), Data Mining, *Morgan Kaufmann Publisher*, 2000, S. Francisco, CA, USA

[5] Sheng Li, Quiang Yang (2001), An Agent System that Integrate Case-Base Reasoning and Active Databases, *International Journal of Knowledge and Information System* (2001) 3:225:251, Spriger-Verlag

[6] M. Mocenico, A. Serrani, M. Sforna, The Dispatcing Organisation of the Italian Independent System Operator, www.grtn.it