

Editoriale

Sistemi Autonomici

È con grande piacere che scrivo queste righe di introduzione del numero speciale di Mondo Digitale dedicato alla evoluzione verso sistemi autonomi simbiotici.

Questo tema è stato al centro delle mie attività negli ultimi due anni in cui ho guidato una iniziativa del Future Direction Committee, FDC, dell'IEEE proprio su questa tematica.

Il FDC ha per obiettivo l'esplorazione di temi in cui l'innovazione tecnologica fa da propulsore su cui convergono svariate tecnologie richiedenti un approccio multidisciplinare e multi-mercato, travalicando quindi le aree di interesse delle diverse aree dell'IEEE (Societies e OUs). Il risultato delle attività portate avanti dall'FDC trova poi una naturale prosecuzione o attraverso la formazione di gruppi di interesse oppure attraverso la presa in carico da parte di alcune società dell'IEEE.

Il tema dei Symbiotic Autonomous Systems ha certamente queste caratteristiche di multi disciplinarietà e di trasversalità sia rispetto alle Societies dell'IEEE sia rispetto all'industria e al mercato.

L'iniziativa nei suoi primi due anni ha portato alla definizione di una roadmap concettuale prima e di una analisi dettagliata sulle singole roadmap tecnologiche poi che sono fruibili attraverso due White Paper (si veda l'articolo introduttivo a questo numero). Inoltre ha evidenziato vari aspetti, ed impatti, in diversi mercati e industrie. Altro elemento importante è stata la analisi sulle possibili implicazioni sociali, etiche e regolatorie.

Che le macchine stiano diventando sempre più flessibili ed in grado di operare in modo autonomo è evidente. Dai robot industriali siamo passati alle metropolitane senza guidatore per arrivare ai robot domestici (come l'aspirapolvere). A queste macchine "hard" si affiancano, sempre più, quelle soft, i soft robot, cybots, che operano nel cyberspace. Questa evoluzione è notevole non tanto, e non solo, per le tecnologie che vengono messe in campo ma per i costi scesi al punto da consentire un ingresso nel mass market e per la parallela capacità di interazione di questi sistemi con noi, in modo naturale –seamless come direbbero gli inglesi.

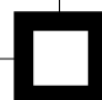
0

1

0

1

0



A questa evoluzione delle macchine sta affiancandosi una non meno rilevante evoluzione delle persone che attraverso la tecnologia aumentano le loro capacità fisiche ed anche quelle cognitive aprendo la porta ad una “human augmentation” che potrebbe sfociare nel transumanesimo. Quest’ultimo è stato considerato nella iniziativa e la conclusione cui si è giunti è di un transumanesimo emergente, derivato dalla simbiosi tra macchine sempre più autonome e intelligenti e persone sempre più aumentate dalle tecnologie e dal contesto tecnologico in cui vivono.

Si è quindi pensato, insieme alla redazione di Mondo Digitale, che voglio ringraziare per il continuo supporto e guida, di preparare questo numero speciale sul tema dei sistemi autonomi simbiotici.

Si parte con un articolo introduttivo da parte mia per passare a considerare due dei pilastri tecnologici su cui si basa questa evoluzione, quello dei Digital Twin, elencati da Gartner tra le tecnologie di maggior impatto nel 2019, a cura di Maurizio Griva, e quello della Intelligenza Artificiale che sta evolvendo verso una Intelligenza comparabile a quella umana (AGI – Artificial General Intelligence) e che nelle prossime decadi verosimilmente la supererà (ASI – Artificial Super Intelligence) a cura di Stuart Dambrot, studioso di questi temi.

L’intelligenza artificiale opera su dati, sulla possibilità di rilevare il mondo esterno e questo passa attraverso sensoristica e comunicazione, aspetti che sono trattati nell’articolo seguente da Massimo Donelli della Università di Trento andando a svolgere interessanti considerazioni anche sulla trasformazione delle stesse reti di comunicazione in sistemi autonomi auto-configuranti con particolare riferimento alle reti radio che saranno impiegate nel 5G.

Il mondo dei sistemi autonomi, proprio per questa autonomia, richiede una riflessione su quanto ci si possa “fidare” di loro. Un tema decisamente nuovo con cui è necessario confrontarsi e a questo è dedicato l’articolo di John Budenske che con tratto leggero affronta questo tema che leggero non è affatto.

In diretta prosecuzione sul tema della “fiducia” Antonio Crespo ci offre una visione a 360° sullo scenario che si apre con i velivoli autonomi e Aman Singh propone di utilizzare sistemi autonomi per analizzare in modo dinamico rischi nel cyberspace e prendere relative contromisure: sistemi autonomi per contrastare sistemi autonomi. Inoltre nel suo articolo Aman spiega come questo approccio basato su sistemi autonomi possa trovare applicazione anche per contrastare tentativi di disinformazione, fake news.

L’articolo successivo, di Francesco Flammini, si focalizza sugli aspetti di sicurezza nella interazione tra persone e macchine, quindi all’interno di sistemi simbiotici, una nuova area che è bene coltivare fin da oggi, facendo notare come trovi applicazione anche nella gestione della sicurezza nel settore IoT.

Si passa quindi ad esaminare alcuni settori “verticali” di applicazione di sistemi autonomi simbiotici, a partire con l’articolo di Domenico Presenza che si focalizza su applicazione di tali sistemi nell’Assisted Living tramite Robot

domestici. Prosegue Paolo Traverso presentando alcuni servizi ora in fase di esercizio sperimentale nella Provincia di Trento facendo quindi emergere gli aspetti di rilascio su di un territorio e l'impatto sulle persone.

L'impatto sulle persone è il focus su cui si chiude questo numero con due articoli di notevole interesse anche per un pubblico allargato. Nel primo Witold Kinsner affronta il tema del cambio di paradigma nel mondo dell'Istruzione (education), nel secondo Derrick de Kerckhove mette in luce l'enorme cambiamento che questi sistemi porteranno a livello sociale, un cambiamento che secondo Derrick ha un parallelo nel cambiamento prodotto dal Rinascimento. Entrambi ritengono che saranno i Digital Twins l'elemento di rottura rispetto al passato, fornendo un legame tra atomi e bit in quella che oggi viene chiamata Trasformazione Digitale. Buona lettura!

Roberto Saracco

Sistemi Autonomi Simbiotici

Roberto Saracco

Sommario

L'automazione sta per fare un salto di qualità: grazie alle evoluzioni tecnologiche e a spinte di mercato l'automazione permetterà non solo di avere macchine sempre più performanti ma di aumentare anche le nostre capacità di esseri umani, sia quelle fisiche sia quelle cognitive. In prospettiva ci si immagina una convergenza tra il mondo –aumentato- delle macchine e quello –aumentato- delle persone che porta ad una vera e propria simbiosi tra uomo e macchine, simbiosi, peraltro, di cui già oggi si vedono i primi segnali.

In questo articolo si presentano gli studi effettuati in ambito di una iniziativa del Future Direction Committee dell'IEEE sui Sistemi Autonomi Simbiotici. Molti dei concetti delineati saranno poi estesi e istanziati a casi specifici trattati negli articoli seguenti.

Keywords: Autonomous Systems, Digital Twins, Human-Machine symbioses, IEEE, Transhumanisms, Brain-Computer Interfaces, Future of Education

Gli esseri umani, come peraltro tutti gli esseri viventi, vivono in simbiosi con il loro ecosistema. Alcuni vivono in una simbiosi stretta, cioè al di fuori di questa non riuscirebbero a vivere (è il caso nostro con il microbioma che popola il nostro intestino e che ci consente di nutrirci...), altri in una simbiosi più debole in cui vi è un vantaggio nello stare insieme ma si può stare anche senza l'un l'altro.

Gli uomini, a differenza di altre forme viventi, hanno, nel tempo, creato un insieme di strumenti e di strutture con cui vivono in "simbiosi", spesso in simbiosi stretta al punto da non poterne fare a meno. Pensiamo alle strutture che ci forniscono energia: senza di queste la vita che viviamo ogni giorno sarebbe difficile e nel giro di poco tempo la vita stessa diventerebbe impossibile. Certo, si potrebbe fare il Robinson Crusoe, ma non potremmo farlo tutti! Robinson aveva tutta un'isola a disposizione, le centinaia di migliaia di persone che popolano una città no: senza infrastruttura energetica che alimenti le catene distributive si troverebbero senza nulla da mangiare, né bere... Non parliamo

poi del sistema sanitario che si fermerebbe, le comunicazioni che sarebbero limitate allo spazio locale... e così via.

Negli ultimi 10 anni abbiamo iniziato a creare strumenti con un crescente grado di autonomia, autonomia che si svilupperà ulteriormente nella prossima decade. Questi strumenti, per questa autonomia ma certamente anche per le loro caratteristiche di flessibilità ed adattabilità, diventano dei sistemi veri e propri, autonomi ed in grado di interagire con il loro ambiente, adeguando il loro comportamento. Questo è reso possibile dalla capacità di realizzare sistemi complessi che hanno una percezione del contesto in cui si trovano ad operare, contesto in cui, sempre più spesso, ci siamo noi.

Siamo quindi di fronte a sistemi autonomi che in qualche modo hanno caratteristiche funzionalmente simili alle nostre –incluso capacità di intelligenza-, al punto da poter entrare in competizione con noi. Non è di oggi la preoccupazione di vedere robot, e sistemi software prendere il nostro posto in svariate (e crescenti) attività.

Nella prossima decade assisteremo a due direttrici di evoluzione (Figura 1). Da un lato noi, sempre più in grado di svolgere compiti complessi e gravosi grazie alla disponibilità di strumenti sempre più sofisticati, dall'altro loro, sistemi autonomi, con una crescente intelligenza in grado di affrontare ed operare in contesti sempre più complessi e variegati.

Queste due direttrici potrebbero trovare un punto di incontro in quella che chiamiamo la simbiosi tra sistemi autonomi, la simbiosi tra uomo e macchina.

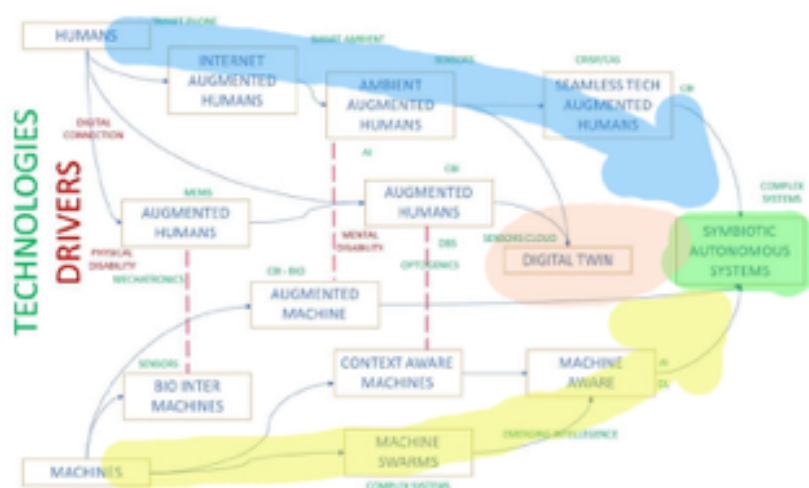


Figura 1

L'evoluzione tecnologica porta ad un incremento delle capacità delle persone da un lato e dall'altro all'incremento delle capacità delle macchine. Questo incremento è anche favorito dalla interazione tra uomo e macchina in cui l'uomo si avvantaggia della macchina e la macchina dell'uomo. In prospettiva questa interazione si rafforzerà diventando usuale e come tale non percepita portando ad una simbiosi tra uomo e macchina.

Esaminando la figura 1, sviluppata nel contesto della Iniziativa Symbiotic Autonomous Systems dell'IEEE, si nota come gli esseri umani stanno aumentando le loro capacità utilizzando strumenti, artefatti che sfruttano l'evoluzione tecnologica.

Internet -il web-, che oggi gran parte delle persone utilizza attraverso il computer e sempre più attraverso lo smartphone, fornisce un aumento alle nostre capacità di memoria (con pochi click riusciamo a sapere tutto – compreso tutto di tutti e questo può essere un problema) ma anche alle nostre capacità intellettive, si pensi alla capacità di calcolo offerta da applicazioni sul web o alla possibilità di capire altre lingue...

La possibilità crescente di avere interfacce trasparenti al web, quali quelle basate su augmented reality, l'utilizzo di accessi come televisori (che stanno diventando sottili come fogli, arrotolabili e sovrapponibili alle pareti come tappezzeria), oggetti che comunicano in linguaggio naturale (si pensi ad Alexa, Siri, ...) portano ad una naturalezza della interazione che annulla la distanza tra una persona e il suo ambiente. Nelle prossime decenni realtà aumentata e virtuale saranno mediate attraverso wearable e implants che annulleranno completamente la percezione di separazione tra noi e gli oggetti spingendo verso una simbiosi funzionale.

A questa evoluzione si affianca l'evoluzione delle protesi sviluppate per supplire a deficit fisici, quali quelli conseguenti alla perdita di un arto o alla perdita di funzionamento a seguito di problemi neurologici (ad esempio ictus) che sta portando a protesi non solo efficaci, in grado di colmare il deficit, ma anche in grado di aumentare le capacità funzionali oltre la norma (alcuni atleti paralimpici stanno diventando in grado di avere prestazioni migliori di quelli normodotati grazie a protesi).

Alle protesi per compensare deficit fisici si stanno affiancando protesi per compensare deficit mentali. Queste si svilupperanno nella prossima decade e se oggi sono focalizzate sui deficit (ad esempio perdita della memoria nell'Alzheimer) nelle prossime decenni saranno in grado di aumentare le capacità mentali, anche attraverso una simbiosi con il cyberspace attraverso interfacce BCI (Brain Computer Interface) che diventeranno invisibili.

Questa fusione tra persona e cyberspace diventa mediata dai Digital Twins, una tecnologia che sta avendo un forte sviluppo da metà di questa decade nel mondo industriale (mediando oggetti/prodotti/processi), che diventeranno il collante della simbiosi tra uomo e macchina.

Nella parte inferiore della figura 1 vediamo che sul versante "macchine" (dove con macchina viene intesa sia la macchina vera e propria, un oggetto che ha un insieme di funzionalità e interagisce con l'ambiente, sia una applicazione software) si assiste ad una progressiva evoluzione che aumenta le funzionalità della macchina anche attraverso l'uso di interazioni con sistemi viventi (bio-interacting machines) e attraverso la collaborazione tra macchine (machine swarms).

Questo porta ad una nuova generazione di macchine che sono molto più flessibili ed autonome in quanto in grado di percepire l'ambiente in cui operano

(context aware machines) e di comportarsi in modo intelligente (machine aware). Questa evoluzione sfocia nella capacità delle macchine di integrarsi con il loro ambiente (ambiente in cui ci siamo anche noi) e in ultima analisi ad una simbiosi con questo (o parte di questo).

L'evoluzione da un lato delle persone e dall'altro delle macchine porta quindi, nel lungo termine, 2040-2050, ad una simbiosi tra uomo e macchina. Il percorso sarà graduale e alcuni effetti sono già percepibili oggi e diventeranno sempre più marcati nelle prossime decadi.

La Iniziativa IEEE su Symbiotic Autonomous Systems - SAS

Nel 2017 il Future Direction Committee dell'IEEE ha lanciato una iniziativa su SAS partendo proprio dalla constatazione che il continuo progresso nella digitalizzazione ed elaborazione dei dati sta portando ad una trasformazione da un mondo di artefatti inanimati ad un mondo di sistemi autonomi con una crescente intelligenza. Questi sistemi stanno acquisendo una capacità di interazione e di consapevolezza del contesto in cui operano tale da farli diventare da un lato sempre più indipendenti e dall'altro sempre più collaborativi sia tra loro sia nei rapporti con persone.

La Iniziativa SAS si pone l'obiettivo di generare un consenso sul modo più efficace per sviluppare una relazione simbiotica tra sistemi autonomi, consapevole di come questo possa avere un impatto rivoluzionario sulla società nelle prossime decadi. L'iniziativa affronta sia gli aspetti tecnologici che abilitano l'evoluzione verso una simbiosi tra artefatti e persone, sia gli aspetti economici, sociali ed etici che questa evoluzione comporta per indirizzarla verso una crescita a favore delle persone e della società in generale.

Nei sui primi due anni l'iniziativa ha aggregato una platea di oltre 600 persone, in prevalenza con interessi e background tecnologico ma anche operanti in settori economici, etici legali e sociali (ELS) che attraverso incontri e gruppi di lavoro hanno identificato diversi aspetti che sono stati distillati in due White Papers (2017 e 2018), il primo sostanzialmente centrato sulle tecnologie abilitanti, il secondo sulle prospettive di evoluzione tecnologica, inclusi i Digital Twins e l'Intelligenza Artificiale nelle sue diverse forme, e sul loro impatto sul versante economico (nuove catene del valore, posti di lavoro), sociale (democrazia e datacracy, istruzione) etico (social ethics vs machine ethics, mind upload e consapevolezza) e legale (responsabilità, privacy, autonomia).

Tecnologie e roadmap

L'iniziativa ha considerato una pluralità di tecnologie, collegandole ai blocchi identificati nella roadmap (Figura 1). Sono state identificate tre aree principali, l'aumento delle macchine, quello delle persone e la simbiosi tra uomo e macchina. Per ciascuna area è stato creato un diagramma che evidenzia come le diverse tecnologie studiate vadano a contribuire alla loro evoluzione (Figura 2).

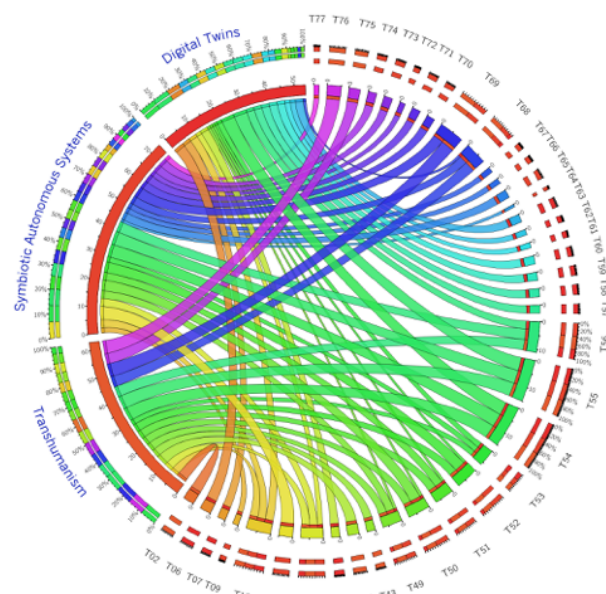
**Figura 2**

Diagramma di contributo delle diverse tecnologie alla base della simbiosi tra uomo e macchina. Le tecnologie sono indicate con le sigle da T02 a T77 (per la lista delle tecnologie si veda il White Paper 2 sezione 3, per quelle indicate in questo diagramma si veda pg. 46 del White Paper II).

- Inoltre, per ciascuna delle tecnologie indicate, e per ciascuna area di applicazione, è stata sviluppata una roadmap andando ad identificare 4 fasi:
- Ricerca a livello accademico con sperimentazioni in laboratorio
- Prime sperimentazioni sul mercato su nicchie in cui i fattori di costo e prestazioni non sono rilevanti,
- Prime applicazioni sul mercato che per estendersi hanno bisogno di una forte diminuzione dei costi, con una ricerca guidata dall'industria,
- Penetrazione su mercato con eventuale trasformazione in commodity.

Avendo per riferimento uno scenario a 20-30 anni non era chiaramente possibile focalizzarsi su una specifica tecnologia. Si sono quindi identificate per ogni area (ad esempio quella dei Digital Twin, Figura 3) un insieme di famiglie tecnologiche (quali ad esempio Sensors' Networks) che si compongono di specifiche tecnologie di cui alcune già presenti oggi, altre che saranno identificate e create nei prossimi anni. È evidente come questo approccio, adatto ad uno studio previsionale di lungo termine, sia necessariamente caratterizzato per prevalenza rispetto all'area funzionale selezionata.

Ad esempio nel caso dei Digital Twins le Sensors' Networks sono considerate oggi in una fase di ricerca industriale, pur essendo già disponibili sul mercato

per alcune classi di applicazione ma hanno ancora un carattere di ricerca industriale per la applicazione ai Digital Twins. Inoltre, le Sensors' Networks utilizzano una varietà di tecnologie di cui alcune già mature (tipo Zig Bee, WiFi...) e altre che saranno sviluppate nei prossimi anni (come i sistemi impulsivi a bassissima potenza e che sono al momento oggetto di ricerca accademica). Inoltre, a livello Sensors' Networks vanno anche considerate le diverse architetture che portano ad elaborazioni distribuite, agli edge, ad analisi semantiche e così via.

Si è anche cercato di mettere in evidenza dei punti di correlazione tra famiglie di tecnologie, ad esempio Neuralink (T49) e BrainInternet (T68) saranno correlate nella quinta decade (verso il 2040) in quanto la possibilità di disporre a tale data di nicchie di mercato per impianti a livello cerebrale che rilevano attività neuronali e le convertono in informazioni semantiche darà verosimilmente un impulso allo sviluppo di connessioni tra cervelli e Internet.

Nel II White Paper vengono illustrate le diverse correlazioni, che a volte assumono un valore rafforzativo (la disponibilità di un progresso tecnologico stimola un altro) a volte portano a deprimere, arrestare, un'area di evoluzione (in quanto lo sviluppo dell'una rende superfluo lo sviluppo dell'altra).

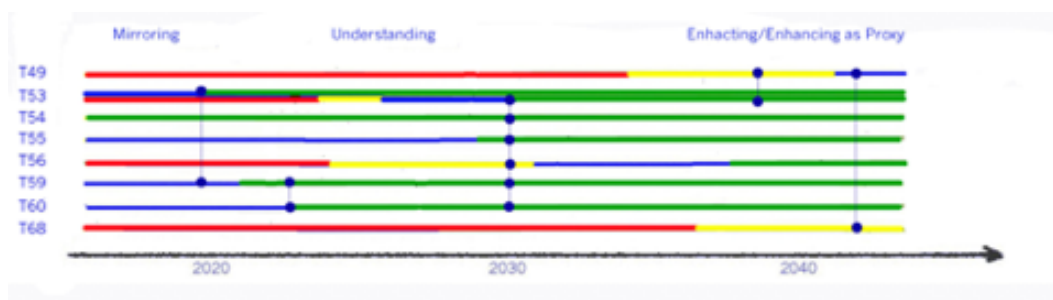


Figura 3

Roadmap delle tecnologie alla base della evoluzione dei Digital Twins. Sono identificati 4 passi evolutivi, dal mirroring di un oggetto alla comprensione semantica dell'oggetto stesso all'aumento dell'oggetto attraverso funzionalità create nel cyberspace e possibilità di agire come proxy dell'oggetto stesso.

In Rosso la fase di sperimentazione di laboratorio, in Giallo quella di prime sperimentazioni sul mercato su nicchie, in Blu le prime applicazioni sul mercato, in Verde la fase di penetrazione su mercato.

- T49: Neuralink
- T53: G5/G6
- T54: Security
- T55: Cybersecurity
- T56: Cyber-Physical Security
- T59: Sensors Networks
- T60: Sensors Cloud
- T68: BrainInternet

Lo studio delle roadmap si presenta molto interessante in quanto permette di avere una comprensione di insieme, mette in luce le correlazioni e gli attori che

guidano l'evoluzione in certe fasi, e in ultima analisi consente di apprezzare le forze che muovono (e promuovono) l'evoluzione fornendo quindi anche elementi per stimolare l'evoluzione andando ad agire sugli elementi chiave. Questo rappresenta un valore soprattutto per le istituzioni (dove investire) e l'industria (quando investire).

Sistemi autonomi simbiotici nella sanità

Il settore della sanità avrà una evoluzione dirompente nelle prossime decadi sotto la spinta di una comprensione del genoma e di come sia possibile influenzare il fenotipo ed il fenotipo esteso.

La comprensione attuale del genoma è ancora limitata: abbiamo etichettato come spazzatura buona parte del DNA in quanto ritenuto vestigia di un processo evolutivo che ha conservato ma non utilizza molto del codice genetico presente nei nostri cromosomi ma ora si sta iniziando a capire che quel codice non è proprio del tutto inutile e non solo, che la attivazione di parti di quel codice (epigenetica) potrebbe portare ad un aumento delle capacità di un individuo.

Se siamo distanti dal sapere tutto del genoma lo siamo ancora di più nel comprendere il legame tra genotipo (il genoma di un individuo) ed il fenotipo. Non sappiamo cioè quale sia l'impatto di una modifica di un pezzo del genoma, di un codone, sul fenotipo (sulle caratteristiche fisiche e cognitive) di una persona.

In questi ultimi due anni si è iniziato ad applicare tecniche di intelligenza artificiale (machine learning, data analytics) per identificare correlazioni tra il genotipo ed il fenotipo. La speranza è di arrivare nel corso della prossima decade ad una comprensione se non completa perlomeno significativa della correlazione, anche grazie al crescente numero di genomi sequenzializzati, che dovrebbe esplodere nel corso della prossima decade grazie all'aumento di velocità di sequenzializzazione e diminuzione del costo sotto la soglia dei 100€.

Abbiamo la tecnologia (CRISPR/Cas 9) per modificare il genoma e nuove tecnologie affinano questa possibilità rendendola sempre più precisa ma senza una comprensione di quale sia l'effetto di una modifica sul fenotipo i progressi sono molto lenti (come avere in mano una pistola precisissima ma avere gli occhi bendati quando la si usa).

Nelle prossime decadi dovrebbe diventare possibile capire come modificare un genoma per ottenere un certo cambiamento nel fenotipo (ammesso che quel cambiamento sia possibile), ad esempio si potrebbero introdurre variazioni al genoma per diventare immuni a batteri e virus, una vaccinazione a priori (questo viene già fatto in qualche misura in agricoltura alterando il genoma della pianta per renderla resistente a certe infezioni). Un cambiamento del genoma potrebbe anche portare alla capacità di nutrirsi con una maggiore varietà di sostanze. Notare come il cambiamento potrebbe essere effettuato sul genoma umano o sul genoma dei batteri che costituiscono il microbioma intestinale (e che oggi ci permettono di digerire certe sostanze). Questo è un punto interessante in quanto mette in luce che in un sistema simbiotico le variazioni possono essere

introdotte potenzialmente su uno qualunque dei suoi componenti generando una variazione complessiva nel sistema simbiotico.

L'evoluzione delle protesi porta ad una evoluzione del fenotipo esteso, amplificando ad esempio la forza e la resistenza di un individuo (vedi sezione su Industria 4.0). L'utilizzo di protesi può estendere la nostra capacità sensoriale, ad esempio quella di percepire lunghezze d'onda nell'infrarosso e ultravioletto, udire suoni al di fuori delle frequenze percepite dal nostro orecchio, rilevare la presenza di molecole/batteri "annusando" l'ambiente.

L'evoluzione del fenotipo esteso attraverso simbiosi con protesi sarà probabilmente quella più seguita anche per i minori risvolti etici (rimane tuttavia l'aspetto del gap tra chi potrà permettersi una protesi e quindi acquisire un vantaggio su chi invece non potrà permetterselo...).

Sistemi autonomi simbiotici nell'industria 4.0

La rivoluzione industriale ha visto l'emergere con prepotenza delle macchine e in parallelo dell'energia per farle funzionare.

Il rapporto tra operaio e macchina è stato ambivalente: da un lato la macchina vissuta come minaccia al lavoro in quanto sostitutiva della forza lavoro umana, dall'altro la macchina che solleva l'uomo dalla fatica. Questa ambivalenza continua oggi con macchine sempre più sofisticate che operano su tutta la catena del valore, non più solo nella fase della produzione. La supply e delivery chain stanno diventando sempre più automatizzate e le macchine in prospettiva diventeranno al tempo stesso strumenti per produrre e prodotti. Vediamo oggi i primi segnali nel settore del software con smartphone che si aggiornano automaticamente, evolvendo nelle prestazioni: sono al tempo stesso prodotti che usiamo e macchine in grado di creare un nuovo prodotto.

La pervasività delle macchine investe quindi tutta la catena del valore, elemento questo che caratterizza l'industria 4.0.

Nei prossimi anni questa pervasività si sovrapporrà ad una simbiosi tra persone e macchine. Un esempio, per capirci, è quello dei lavoratori alla catena di montaggio della Ford che utilizzano a partire dal 2017 l'esoscheletro della Exsoworks, dei robot indossabili che riducono sensibilmente la fatica. L'esoscheletro tramite sensori rileva il movimento delle braccia e interviene con le proprie braccia che affiancano quelle dell'operaio fornendo un supporto che toglie il peso dagli arti, riducendo quindi la fatica.

Nei prossimi anni gli esoscheletri si integreranno sempre più con il corpo dell'operaio fino a diventare una vera e propria estensione che opererà in simbiosi con il corpo umano. L'interazione tra esoscheletro e corpo avverrà tramite un insieme di tecnologie che rilevano i segnali nervosi e anche la mimica posturale e facciale con del software in grado di interpretarli per tradurli in azione. Sono già in fase di sperimentazione in laboratorio dei guanti che estendono il numero di dita della mano (in genere aggiungendo un pollice). Si vede che dopo poco tempo il cervello della persona che li indossa integra questo "dito" aggiuntivo nella mappa del corpo e inizia ad utilizzarlo un modo trasparente (senza che la persona debba pensarci, così come non pensiamo a

come muovere le dita per prendere un uovo dal tavolo).

Questo è un vero e proprio passo verso la simbiosi e l'intero comparto dell'industria 4.0, segnato da una collaborazione seamless tra robot e persone in tutta la catena del valore trarrà beneficio da questa e al tempo stesso darà un forte impulso alla simbiosi tra uomo e macchina.

Parte robotica e essere umano, entrambi sistemi autonomi, convergono creando una nuova entità. Resta ovviamente aperto il tema su quanto questa evoluzione sarà accettata, addirittura auspicata, dai lavoratori e quanto possa venire imposta dai datori di lavoro.

Impatti sull'istruzione

Un aspetto importante, evidenziato dallo studio portato avanti nella iniziativa dell'IEEE, è quello dell'istruzione, sia a livello scolastico/universitario sia a livello professionale (continuous education).

La quantità di conoscenza oggi disponibile ha superato la capacità della singola persona di assimilarla. Non solo. La rapidità con cui la conoscenza cresce, in molti settori dalla medicina alla tecnologia ad esempio, è maggiore della capacità del singolo di imparare quindi la distanza tra le conoscenze possedute e quelle disponibili aumenta.

Per ovviare a questa situazione abbiamo creato una simbiosi con il web: oggi è diventato non solo normale ma anche naturale fare affidamento sulle macchine per attingere alla conoscenza. Quanti di noi sono in grado di ricordare il numero di telefono anche solo degli amici? Non serve più, ovviamente, tanto per telefonare lo si fa attraverso il nome, non più attraverso il numero di telefono. Gli ingegneri progettano la struttura in cemento armato avvalendosi di programmi, non usano più le formule, il medico consulta il web per trovare una corrispondenza tra sintomi e patologia e tra patologia e medicinali (non tutti ma sempre di più e quelli che fanno così sono quelli che dimostrano maggiore professionalità!). Perfino i piloti di linea sono obbligati a seguire le check list piuttosto che fidarsi della propria memoria...

Unanimous AI, una azienda americana, fornisce un servizio ai medici in cui questi possono attingere alla conoscenza di migliaia di loro colleghi integrata ed aumentata dalla conoscenza elaborata da sistemi di machine learning che vanno ad esaminare centinaia di migliaia di articoli di ricerca in medicina.

La Boeing fornisce strumenti di realtà aumentata ai tecnici che devono effettuare manutenzione sugli aerei per assicurarsi che questi dispongano delle conoscenze necessarie ed altrettanto fa General Electric per il personale di manutenzione delle turbine.

La simbiosi della conoscenza, in cui questa è condivisa tra il nostro cervello e il web è ormai una realtà.

Nell'ambito della iniziativa SAS si è sviluppato un modello in cui la conoscenza di una persona può essere consolidata nel Digital Twin della persona. Questo, come il digital twin di una turbina, è un modello che riflette lo stato attuale delle conoscenze (del funzionamento) ma a differenza di quello il digital twin di una

persona deve anche cercare di sapere quello che la persona può avere dimenticato.

L'idea è di utilizzare il digital twin come un mediatore che da un lato sa quello che il corrispondente physical twin (la persona) sa e dall'altro interagisce con applicazioni che mettono in luce quello che occorrerebbe sapere per svolgere una certa attività. Su questa base viene anche determinato il modo migliore per aumentare la conoscenza nel momento in cui questo è richiesto: la formazione diventa personalizzata rispetto sia alla conoscenza pregressa sia a quelle che è necessario acquisire. Queste ultime, inoltre, tengono conto del contesto per cui si valuta se certe conoscenze non possano essere acquisite da altri che operano nel team, ivi compreso robot e agenti software.

L'intermediazione tramite Digital Twin nel settore dell'apprendimento è quello che rende possibile (o perlomeno facilita) una simbiosi nella conoscenza. Non è importante dove la conoscenza risieda ma che questa diventi fruibile nel momento in cui serve.

Si noti anche che l'approccio con Digital Twin apre delle questioni nuove: se la mia conoscenza risiede anche nel mio digital twin, l'azienda per cui lavoro potrebbe decidere di tenersi una copia del mio digital twin nel momento in cui lascio l'azienda per andare in un'altra, o magari potrebbe decidere di trasferire le conoscenze dal mio digital twin a quello di un robot in modo che il robot acquisisca le mie conoscenze...

La simbiosi .. dei Self

Un organismo simbiotico emerge dalla simbiosi tra i suoi componenti ed in una certa misura è rappresentabile dall'emergere di un "self" (che può o meno essere conscio).

Pensiamo a noi stessi. In quanto individuo ho ben chiaro il senso di essere tale. Il mio self è parte integrante dell'essere io. Nel caso delle persone il self (in generale ed in assenza di deficit cognitivi) è anche percepito a livello conscio, sono consapevole che sono io.

Ora immaginiamo di aver perso un arto. Il senso del self rimane integro, sono sempre io, anche se, ovviamente, le mie capacità potranno risultare diminuite dalla perdita dell'arto. Il self, quindi non è legato alla integrità delle parti ma alla identità associata.

Ora pensiamo di supplire alla mancanza dell'arto attraverso una protesi di ultima generazione. Queste protesi sono in grado di interagire in modo trasparente con il mio cervello (attraverso varie tecnologie dall'impianto di chip alla intelligenza artificiale) e il mio cervello grazie alla sua plasticità nel giro di qualche mese incorpora la protesi nella sua mappa (questo lo si è dimostrato in una molteplicità di situazioni, ad esempio nella capacità del cervello di convertire segnali sonori in segnali luminosi tramite sinestesia). A questo punto la protesi diventa parte integrale di me, del mio self. Sono entrato in simbiosi con l'arto artificiale.

Pensiamo ora a delle protesi che mi consentano di capire un'altra lingua. Protesi auditive come le sleep-buds della Bose sono state progettate per essere integrate non solo nell'orecchio ma anche a livello funzionale nel cervello (facilitano il sonno e in alcuni diminuiscono/annullano il tinnito). Una protesi di questo tipo potrebbe consentire nell'arco della prossima decade di ascoltare in italiano una persona che mi sta parlando in cinese. Protesi simmetriche potrebbero consentire al mio interlocutore cinese di convertire il mio italiano in cinese.

Nel tempo queste protesi, proprio perché trasparenti alla percezione, diventerebbero parte integrante di me, parte del mio self. Anche qui abbiamo un caso di simbiosi, di natura aumentativa in quanto sono in grado di fare più cose.

Abbiamo visto come l'evoluzione vada verso la disponibilità di protesi cognitive (la comprensione di un altro linguaggio è probabilmente a cavallo tra il fisico e il cognitivo) e come questo possa essere realizzato (in prospettiva) attraverso la tecnologia dei Digital Twins. Nel momento in cui diventi possibile avere una interazione trasparente con il proprio Digital Twin questo diventerà parte del mio self. In altre parole, attraverso il mio Digital Twin il mio self si estende nel cyberspace. Questa è una estensione di cui posso essere conscio ma per la maggior parte dei casi non lo sarò, sarà semplicemente un dato di fatto che, ad esempio, possa avere informazioni, e sensazioni, consce mediate (generate) dal digital twin sulla base di esperienze che questo vive nel cyberspace.

Una domanda che si pone è se il Digital Twin, da parte sua, sia conscio. Questa è una domanda a cui non è possibile rispondere, almeno oggi. Certamente, se si accettasse un principio simile a quello del test di Turing per l'intelligenza, alcuni potrebbero concludere che in prospettiva il Digital Twin di una persona ha consapevolezza (e analogamente quello di una macchina che sia conscia sarebbe conscio).

Come è possibile che vi siano due self (il mio fisico e quello del mio Digital Twin) co-presenti nel mio self? In effetti la cosa non sarebbe né strana né una novità. Esperimenti negli ultimi anni su come il cervello percepisce il mondo esterno e come fa emergere il senso di sé hanno chiaramente evidenziato come in ogni persona esistano più self spesso in contrapposizione tra loro (Dr Jekyll e Mr Hyde non sono più una creazione immaginifica) che in caso di patologie neurologiche perdono la capacità di integrazione portando all'emergere di personalità diverse.

Queste molteplicità di self che la tecnologia promette di produrre possono trovare, come trovano nel nostro cervello, una sintesi in un super-self dell'elemento simbiotico.

Questi sono temi affascinanti ancora tutti da scoprire che portano con sé anche nuovi aspetti etici.

Aspetti Etici

L'evoluzione verso sistemi simbiotici, in particolare per le relazioni tra i loro componenti e l'emergere di un sistema simbiotico apre la porta su orizzonti ancora tutti da esplorare di cui si iniziano già oggi a vedere alcune implicazioni, basti pensare ai veicoli autonomi, le macchine che guidano da sole.

2019-2034	2035 – 2050	
Sistemi Autonomi Simbiotici visti come "Oggetti"	Etica Sociale (e.g., Perdita posti di lavoro) Etica delle macchine: algoritmi etici Equità negli algoritmi	Trasferire ns valori umani alle macchine Prevalenza degli Agenti Gestire AGI e ASI
Sistemi Autonomi Simbiotici visti come "Esseri"	Stato legale SA Controllo dell'apprendimento Condivisione responsabilità	Stato legale SAS Consapevolezza - Coscienza Trasferimento (upload) della Mente – diritti, responsabilità

Tabella 1
Aspetti etici relativi a sistemi autonomi simbiotici

Oggi siamo già sottoposti al monitoraggio (controllo?) di sistemi elettronici automatici e in alcune nazioni (ad esempio Singapore) questo monitoraggio si traduce anche in una forzatura sui comportamenti con uno spostamento della responsabilità dal singolo (self) ad un ordine sociale che diventa auto-organizzato in cui le scelte su priorità sono prescritte dal sistema (da chi lo controlla). Se questo trend dovesse rinforzarsi occorrerà trasferire i valori etici dal singolo componente al sistema, occorrerà codificare una etica nella Intelligenza Artificiale. Ovviamente questo spostamento dal self al super-self pone il problema di chi sia titolato ad "istruire" il super-self.

Non sarebbe la prima volta in cui la "natura umana" e i suoi valori cambiano: si pensi che qualche secolo fa le strutture sociali erano fortemente gerarchiche e il diritto del principe prevaleva sul diritto dei sudditi, l'immobilismo era un valore positivo, l'innovazione un valore negativo... Queste diversità erano fortemente interiorizzate nei singoli "self" e il processo educativo ne era uno strumento e una conseguenza.

Con la diffusione dei Digital Twins la responsabilità (e anche l'istruzione) potrebbe spostarsi dal self fisico a quello digitale. L'emergere di una simbiosi porterà ad una ridefinizione dei confini del self (e connesse responsabilità)? Si noti come questi elementi inizino ad essere percepiti, come osservavo, quando si considerano le auto a guida autonoma. Alcuni incidenti hanno evidenziato come il guidatore si "fidi" dell'auto al punto da distogliere la sua attenzione. Oggi questo viene considerato una responsabilità del guidatore che ... non avrebbe dovuto. In realtà sappiamo che a livello psicologico (e lo vediamo nei piloti che usano l'ILS) si arriva ad un punto per cui ci si fida della macchina (e in alcuni casi occorre fidarsi!) per cui avviene implicitamente un trasferimento di responsabilità dal self fisico al super-self.

La possibilità di aumentare le proprie capacità fisiche e cognitive attraverso l'epigenetica (prima) e la genetica (poi) potrebbe essere considerata la strada per controbattere l'emergere dell'AGI e evitare l'ASI ma questo, se avvenisse, sarebbe patrimonio di pochi. Il contrasto all'ASI si tradurrebbe quindi in un contrasto tra chi è aumentato e chi no, tra chi partecipa ad un super-self e chi no.

È evidente, anche da queste poche osservazioni come i problemi etici da affrontare saranno molti e nuovi e quanto sia necessario anticiparli per guidare l'evoluzione su direttrici desiderabili.

Conclusioni

L'evoluzione verso sistemi autonomi è inevitabile e porta con sé una prevedibile simbiosi tra uomo e macchina, che già esiste, come osservato all'inizio, ma che è destinata ad espandersi nelle prossime decadi.

Il progresso tecnologico, in molte aree, tenderà a convergere verso una evoluzione sistemica che a sua volta è destinata ad accelerare l'evoluzione delle singole componenti tecnologiche come osservato nelle roadmap.

Oggi i sistemi autonomi sono relativamente pochi e operano in settori indipendenti. Nelle prossime decadi il loro numero aumenterà notevolmente e si inizierà ad osservare una progressiva interazione tra questi sistemi, ad esempio in ambito urbano. Questa integrazione si tradurrà, nel tempo, in una simbiosi che coinvolgerà le persone.

Alcuni elementi di questa simbiosi anticiperanno l'evoluzione in alcuni settori, tra questi quello della salute, dell'Industria 4.0 e dell'istruzione.

Bibliografia

SAS White Paper 1:

<https://symbiotic-autonomous-systems.ieee.org/images/files/pdf/sas-white-paper-final-nov12-2017.pdf>

SAS White Paper 2:

<https://symbiotic-autonomous-systems.ieee.org/images/files/pdf/SAS-WP-II-2018-Finalv3.2.pdf>

Televisori ultrasottili CES 2019:

<https://www.theverge.com/2019/1/7/18171013/lg-rollable-tv-oled-4k-tv-features-photos-video-release-ces-2019>

Intelligenza Artificiale per collegare genotipo e fenotipo:

<https://www.mdpi.com/2073-4425/10/1/18>

Influenza del microbioma sulla capacità di digerire sostanze:

<https://www.ncbi.nlm.nih.gov/books/NBK154098/>

Utilizzo esoscheletri nella catena di montaggio Ford:

<https://www.engadget.com/2018/08/07/ford-exoskeletons-eksovest/>

Servizi di accesso a conoscenza in medicina:

<https://unanimous.ai>

Biografia

Roberto Saracco si è innamorato della tecnologia e dei suoi impatti molto tempo fa. Ha un background in matematica, fisica e computer science. Fino ad aprile 2017 ha diretto il nodo italiano di EIT Digital ed è quindi stato responsabile dell'Industrial Doctoral School dell'EIT Digital fino a settembre 2018. In precedenza, fino a dicembre 2011 è stato direttore del Telecom Italia Future Centre a Venezia, studiando le relazioni tra evoluzione tecnologica, economia e società. Intorno al 2000 ha condotto un progetto della Banca Mondiale, Infodev, per stimolare la crescita imprenditoriale in America Latina.



E' senior member dell'IEEE dove presiede l'Industry Advisory Board nell'ambito del Future Directions Committee ed è responsabile della Symbiotic Autonomous Systems Initiative e co-responsabile della iniziativa su Digital Reality e Digital Transformation. Insegna un corso al Master su Previsioni Tecnologiche e Impatti di Mercato alla Università di Trento.

Ha pubblicato oltre 200 articoli in riviste e giornali e 14 libri.

Cura un blog giornaliero <http://sites.ieee.org/futuredirections/category/blog/>, commentando l'evoluzione tecnologia e l'impatto su società e mercato in vari settori.

Email: roberto.saracco@gmail.com

Digital Twin

Maurizio Griva

Sommario

I Digital Twin sono una tecnologia emergente che promette di cambiare lo scenario in molti settori costituendo un elemento importante nell'abilitare la trasformazione digitale. Da un lato i Digital Twin sono una realtà in alcuni, importanti, ambiti industriali, dall'altro tuttavia richiedono come tutte le tecnologie emergenti di raggiungere uno stato di maturazione che ne renda possibile una applicazione generalizzata. In questo articolo si esaminano gli aspetti concettuali per poi calarli nelle realtà industriali in cui questi potrebbero essere adottati analizzando i temi da affrontare per renderli efficaci, basandosi sulla esperienza di Reply nel proporre soluzioni innovative.

Keywords: Digital twin, autonomous systems, AI, IoT, predictive maintenance, industry 4.0

La quantità incredibile di dati, applicazioni, informazioni e conoscenze che continua ad accumularsi sul web sta progressivamente spostando la capacità di imparare come caratteristica umana verso una crescente possibilità di aumentare la conoscenza accedendo al web.

Ci stiamo muovendo, lentamente ma inevitabilmente, da un focus sulla conoscenza personale, accumulata e organizzata individualmente grazie a percorsi formativi unici, a un focus sull'accesso alla conoscenza. Ciò ha profonde implicazioni per l'industria, l'istruzione e la società. Le aziende stanno diventando sempre più consapevoli della necessità di un aggiornamento continuo del loro "know-how aziendale" e mentre in passato questo è stato distribuito tra le loro forze di lavoro oggi è in parte nel cyberspazio, considerato quindi come uno strumento di accesso alla conoscenza stessa. Il vantaggio competitivo di un'azienda consiste perciò nel disporre di una forza lavoro smart in grado di accedere e applicare tale conoscenza e di far sì che la conoscenza sia posseduta esclusivamente - almeno per un certo periodo di tempo - dall'azienda.

Il tutto è integrato in un processo rapidissimo che vede l'interazione tra esseri umani, robot, software e dati in un continuum cognitivo e produttivo.

Questa trasformazione da un mondo dominato da asset fisici ad uno largamente virtuale, nel senso di tangibile, solo attraverso un'estensione dei sensi umani agli spazi digitali, viene spesso denominata come Digital Transformation, almeno in alcune sue componenti.

La Digital Transformation smaterializza gli elementi del mondo fisico e li rappresenta in un mondo di dati e informazioni digitali. Questa smaterializzazione offre vantaggi in quanto i bit possono essere spostati, copiati, aggregati e rielaborati a costi infinitesimi ed a velocità prossime a quelle della luce. I Digital Twin sono una tecnologia a supporto di questa smaterializzazione, replicando in informazioni, dati e programmi gli oggetti fisici composti da atomi.

I Digital Twin sono l'evoluzione dei simulatori digitali ed in quanto tali partono dal concetto di modello digitale di un'entità, fisica o virtuale, semplice o complessa e di sistema di sistemi.

Un modello digitale di un'entità fisica è una replica virtuale di tale entità o sistema descritta con un linguaggio formale adatto ad essere elaborato da un computer. Ogni modello digitale ha una sua intrinseca accuratezza di rappresentazione del comportamento, in termini di input e output, delle entità associate, e può reagire pertanto a stimoli esterni in forma, ad esempio di istruzione di controllo, segnali provenienti da sistemi reali o ad altri componenti, reali o virtuali anch'essi, che con esso interagiscono.

L'avvento dei Digital Twins sta anche ridisegnando la conoscenza e il panorama cognitivo, in cui un'azienda può rivendicare la conoscenza del Digital Twin della propria forza lavoro (ad esempio, la conoscenza dei brevetti o conoscenze specialistiche ad alto valore) e continuare a usarla una volta che il dipendente lascia la società.

Un Digital Twin è la replica digitale di un asset, ma anche potenzialmente di un concetto astratto, di un'idea, di un processo, di una tendenza, di un'espressione artistica. Ha la capacità di rappresentare nel mondo digitale un'entità, ed è valido ed efficace in questa mimica quanto è efficace e valido il modello sottostante.

La sfida più complessa del momento è proprio quella di costruire un modello "azionabile" dell'entità che si vuole digitalizzare, ai fini degli obiettivi che si vogliono ottenere con il modello stesso. Il modello, ad esempio, può essere simulato offline rispetto al gemello fisico che rappresenta, e i risultati ottenuti con il modello vengono confrontati con gli output reali, allo scopo di determinare la bontà del modello ma anche di scoprire differenze di comportamento che possono indicare deviazioni dal funzionamento corretto previsto per il sistema nel mondo reale.

Si pone quindi una questione di interazione con il Digital Twin, che è mediata dalle interfacce uomo-macchina convenzionali come applicazioni mobile, schermi di computer, superfici touch, controllo vocale, social network e poco altro al momento. Tuttavia vediamo apparire i segnali prodromici di una futura interazione simbiotica, dove i Digital Twin saranno gli elementi abilitanti in questa evoluzione

portando oggetti fisici, compresi gli umani, ad abitare e interagire nel cyberspazio. Una vera simbiosi è chiaramente lontana, forse oltre l'orizzonte attuale, ma molti ritengono questa sia la destinazione finale immaginabile ora.

Rimanendo con i piedi per terra, i Digital Twin sono già qui oggi, ed hanno applicazioni reali.

Il Digital Twin è spesso considerato un simulatore, ma in realtà va oltre la simulazione stessa perché può essere eseguito in parallelo ad un sistema reale e soprattutto in tempo reale e diventa un vero e proprio riferimento digitale vivente, continuamente alimentato dagli stessi dati che il Sistema reale riceve e continuamente validato dagli output che produce il sistema stesso.

Ogni deviazione è sintomo di un'anomalia del Sistema reale rispetto al suo gemello "perfetto" e quindi determina un allarme o un warning da considerare per migliorarne l'affidabilità, la resilienza, le performance e la sicurezza. Ma non è tutto: il gemello digitale può essere eseguito a velocità superiore del reale, fornendo una vista sul futuro, più o meno prossimo a seconda di quanto preciso e dettagliato è stato realizzato il modello stesso.

Le simulazioni predittive dalla replica virtuale di un prodotto reale nel tempo possono abbassare inoltre i costi dei reparti aziendali che si occupano di ricerca, grazie alle condizioni operative ottimali durante le quali testare i consumi e alle caratteristiche quantitative e qualitative del Sistema, processo o prodotto, cui è associato il Digital Twin; al tempo stesso, aumentano le possibilità di effettuare modifiche anche minime ai processi e di valutarne gli impatti immediati, a breve e lungo termine (analisi "what-if"). La simulazione offline dei robot consente, ad esempio, di controllare in modo esatto l'evoluzione del pezzo lavorato dalla macchina, potendo minimizzare completamente la comparsa di anomalie di realizzazione e assicurando così un risultato quanto più vicino alla perfezione.

Questa capacità predittiva è estremamente preziosa per molti settori applicativi dove i Digital Twin sono già una realtà, come ad esempio l'industria manifatturiera, il commercio, il marketing, il social network, la sanità. Ma altri settori che richiedono la validazione di schemi complessi a valle di una progettazione estesa e spesso costosa quali la pianificazione urbana, la logistica, i trasporti, l'edilizia stanno iniziando a valutare l'applicazione dei Digital Twin quale elemento di trasformazione digitale dei loro processi.

I cambiamenti nell'industria manifatturiera negli ultimi dieci anni sono stati sbalorditivi, ed hanno spinto le aziende verso percorsi di riduzione dei costi, aumento della produttività dei processi ma anche della qualità di prodotti e semilavorati, maggior flessibilità nel rispondere alle esigenze dei loro clienti e l'adozione ad un approccio più globale ai mercati.

Le aziende manifatturiere oggi usano la tecnologia per supportare molti aspetti delle loro operations. La maggior parte dei processi produttivi richiede un avanzato supporto tecnologico in tutte le fasi, dalla progettazione, alla realizzazione, dal controllo all'ottimizzazione alla sicurezza.

Mentre alcuni produttori monitorano da vicino le innovazioni tecnologiche e adottano le migliori disponibili per le loro attività non appena queste diventano

sufficientemente mature, altri continuano a tollerare sistemi legacy che svolgono un determinato lavoro ma limitano flessibilità e innovazione nel ciclo di vita del prodotto.

Le nuove tecnologie oggi si chiamano nel loro complesso con il nome di Industria 4.0 e riassumono l'introduzione dei progressi tecnologici offerti in Internet of Things, cloud e big data nell'industria contemporanea.

Le aziende italiane sono però particolarmente restie a questi cambiamenti, per motivi prettamente culturali, e si percepisce un evidente disallineamento tra obiettivi di miglioramento e "readiness" tecnologica.

I Digital Twin non fanno differenza: la preparazione dei sistemi interni verso la realizzazione del "gemello digitale" non è uniforme nelle aziende e sebbene i sistemi legacy all'interno di un'azienda possano contenere grandi quantità di dati estremamente utili per la costruzione del modello, la separazione delle sorgenti dei dati tra silos organizzativi ne annulla la potenzialità.

L'adozione trasversale di una piattaforma che consenta la creazione organica di Digital Twin di prodotto e di processo può interrompere l'isolamento reciproco tra i silos aziendali. L'obiettivo di questi twin di prodotto, in particolare, dovrebbe essere quello di creare un'esperienza unificata di prodotto indistinguibile da quella reale. Il Digital Twin "aziendale" può quindi costituire un lubrificante per la collaborazione tra engineering e manufacturing, tra design e marketing di uno specifico prodotto o servizio.

Come, quindi, costruire un Digital Twin per la propria azienda?

La prima decisione fondamentale è comprendere cosa deve essere rappresentato dal gemello digitale: un prodotto connesso, ad esempio, o un processo o una conoscenza espressa dinamicamente, cioè nella sua azione all'interno dell'azienda stessa, come ad esempio la capacità di riconciliazione delle fatture in ingresso. I prodotti connessi sono nativamente dotati, tramite le tecnologie IoT che implementano, a fornire dati in tempo reale sulle proprie performance e sullo stato di funzionamento, e come tali potranno essere collegati direttamente al Digital Twin che li rappresenta nel mondo digitale. Parallelamente, un processo tecnologico, ad esempio l'assemblaggio di un prodotto di elettronica di consumo, può fornire in tempo reale dati sull'esecuzione del processo stesso, sulla disponibilità e allarmistica delle macchine utensili ma anche informazioni sulla pianificazione della produzione tramite il MES, sulla disponibilità di materie prime e semilavorati dal Warehouse Management System (WMS) e dati di qualità sul prodotto in assemblaggio grazie all'interfacciamento diretto dei sistemi di test automatico, di machine vision e di cognitive computing.

Il passaggio successivo è quello di identificare un ambiente digitale dove fare "vivere" il proprio gemello digitale: una Digital Twin Platform. Al momento, sebbene alcuni prodotti software in cloud si siano proposti come Digital Twin Platforms, siamo ancora alle prime fasi di sviluppo di tali soluzioni informatiche ragion per cui le variabili in gioco nella loro scelta, sviluppo e personalizzazione sono ancora moltissime. E' infatti ancora poco chiaro quali possano essere le tecnologie allo stato dell'arte richieste per costruire e distribuire gemelli digitali,

come gestire migliaia di gemelli digitali o migliaia di diversi tipi di gemelli digitali e come si integrino i gemelli digitali con altri sistemi aziendali.

Per le organizzazioni che vogliono scalare una strategia basata sui Digital Twin, la scelta da fare fin da subito è quella di implementare una piattaforma per Digital Twins.

Le principali caratteristiche da considerare in una scelta come questa sono:

- **Connettività verso un'unica sorgente di dati affidabile:**

Si suppone che un gemello digitale sia una replica esatta di un asset e i dati provengano da questo asset o processo. Tuttavia, la manutenzione dell'asset, o un cambiamento nel processo, possono modificare lo stato fisico e logico dell'oggetto fisico. Ad esempio, una turbina eolica potrebbe ricevere una parte di ricambio o una versione diversa del firmware installato; altre pale eoliche appartenenti allo stesso parco potrebbero non avere gli stessi aggiornamenti. Una piattaforma per Digital Twin deve essere in grado di aggiornare e fornire lo stato esatto per ogni singolo gemello digitale, definendo rigidamente quale sia la "source of truth" dei dati in arrivo.

- **Gestire il ciclo di vita del Digital Twin**

Come progettare, costruire, testare, implementare e mantenere un gemello digitale e il suo master digitale, cioè la sua implementazione di riferimento? I gemelli digitali sono le istanze di qualcosa chiamato master digitale o Digital Thread. Ad esempio, ciascuna turbina eolica sarà basata su un master digitale comune che rappresenta gli schemi meccanici ed elettrici, la distinta base dei componenti, le versioni del software e altri accessori, fisici o virtuali, utilizzati per creare ciascuna turbina eolica. Una piattaforma per Digital Twin deve fornire degli strumenti semplici per riunire tutte queste informazioni per creare il master digitale e quindi gestire eventuali modifiche al master digitale, così come la possibilità di istanziare rapidamente un gemello digitale a partire da un master, e di derivare nuovi master per modifica rispetto agli esistenti. Altri strumenti software devono successivamente essere disponibili per testare, implementare e gestire ciascun gemello digitale basato su uno specifico master digitale. Gli strumenti devono anche essere in grado di gestire centinaia di master digitali e migliaia di gemelli digitali in una stessa piattaforma.

- **Open APIs**

Un gemello digitale ben definito diventa l'interfaccia e il punto di integrazione per una soluzione IoT industriale. Questo aspetto dovrebbe essere considerato fin dal primo momento in ogni progetto di industrial IoT, in quanto il monitoraggio via dashboard che spesso viene realizzato come prima funzionalità di base dovrà necessariamente vedere in parallelo un Digital Twin in un prossimo futuro. Una piattaforma per Digital Twin deve fornire un set di API (Application Programming Interface) aperte per consentire a qualsiasi sistema software di interagire con il gemello digitale. Ad esempio, i servizi di machine learning e di analisi dovrebbero essere in

grado di interagire con un gemello digitale attraverso un'API. Un'organizzazione dovrebbe essere in grado di integrare il proprio gemello digitale in altri sistemi aziendali, come i sistemi ERP, MES, IIOT platforms o SCM. Un'API di un Digital Twin dovrebbe quindi abilitare il concetto di device-as-a-service, dove il dispositivo reale e il suo gemello digitale sono istanziati contemporaneamente e sono elaborati all'interno della stessa piattaforma.

- **Gestione integrate degli eventi**

In una piattaforma per Digital Twin deve essere possibile impostare eventi e disegnare processi aziendali da eseguire all'interno dei Digital Twin stessi. Ad esempio, eventi per pianificare le chiamate di manutenzione in base a dati in tempo reale e una visualizzazione accurata dello stato corrente di manutenzione dell'asset fisico, o alternative al processo di manufacturing quando si saturano buffer di linea o si verificano micro-fermate che alterano la qualità del prodotto.

- **Visualizzazione**

Le aziende dovrebbero essere in grado di sfruttare la propria piattaforma per Digital Twins per creare visualizzazioni di dati, dashboard e statistiche sui dati storici approfondite in diretta dal gemello digitale. I dati in tempo reale dovrebbero essere collegati al master digitale per consentire il drill-down nei documenti di progettazione, per i prodotti connessi, o altri componenti del master digitale per i processi.

- **Analysis & Analytics**

La piattaforma per Digital Twin deve essere in grado di fornire strumenti evoluti per l'analisi dei dati raccolti nel passato e presente (analysis) e di quelli di forecast e trending (analytics). Qui le tecniche tradizionali di analisi statistica e di quelle fornite dal Machine Learning e Deep Learning si fondono insieme senza soluzione di continuità. Il Digital Twin per sua stessa natura è a-temporale, per cui può essere "accelerato" o stimolato con migliaia di scenari differenti e alternativi per estrarne eventuali pattern di funzionamento non evidenti nel modello statico. Anche questi dati devono essere propriamente visualizzati in modo da risultare immediatamente utilizzabili da tutti gli stakeholder aziendali.

- **Gestione dei dati di progettazione tra i partner della supply chain**

Per realizzare il vero valore del gemello digitale, l'azienda deve avere un approccio globale per raccogliere, gestire e manipolare i dati digitali del prodotto. Una stretta integrazione tra partner e fornitori è essenziale per garantire che il gemello digitale mantenga accuratamente le configurazioni digitali e fisiche. All'evolversi del prodotto fisico si deve evolvere anche parallelamente il suo Digital Twin, integrando i dati di progettazione con quelli di sviluppo del gemello digitale condiviso tra partner e fornitori sulla supply chain.

Le principali piattaforme di cloud computing forniscono già strumenti avanzati per la progettazione e l'esecuzione di tali modelli, ma ancora siamo in una fase acerba dello sviluppo di veri e propri ambienti "domestici" per i gemelli digitali,

tendenzialmente ospitati in cloud data center in modalità Platform as a Service (PaaS).

Simulazioni su larga scala, cioè a livello di fabbrica, di supply chain o di ecosistema di prodotti connessi, richiedono che ogni oggetto sia dotato di Digital Twin.

Alcune grandi aziende sono pioniere nello sviluppo di Digital Twin, tra tutte sicuramente va menzionata General Electric, che li utilizza sistematicamente per identificare comportamenti anomali delle proprie turbine, facendo “nascere” contemporaneamente la turbina ed il suo gemello digitale e monitorando l'evoluzione dei parametri vitali nel tempo di entrambi allo scopo di calcolarne con accuratezza la vita funzionale residua. GE ha sviluppato anche un Digital Twin Framework che consente di creare e ospitare Digital Twin multipli e di farli interagire tra di loro mentre reagiscono ad input esterni sotto forma di flussi informative provenienti da sensori, modelli analitici e dati storici.

Altre piattaforme di derivazione IoT hanno recentemente dimostrato la loro “readiness” verso i Digital Twin. PTC ad esempio con la piattaforma ThingWorks ha dichiarato di poter ospitare Digital Twin orientati a qualunque connected product. IBM con la sua applicazione cloud Watson IoT offre connettori e capacità di analisi machine learning per la creazione di Digital Twin per vari settori di business. SIEMENS con il proprio PLM software ha lanciato una piattaforma integrata con i dispositivi IoT e basata sul nuovo prodotto MindSphere, offrendo una data chain integrata che va dal sensore al channel bus, al PLC fino al gateway intelligente, al dispositivo di Edge Computing e al proprio cloud. Tra le interessanti iniziative open source va riportato Eclipse Ditto, una vera e propria shell per la creazione e l'esecuzione di Digital Twins interconnessi nativamente con il mondo fisico dei dispositivi IoT tramite la virtualizzazione in “things”.

Reply come system integrator innovativo e da molti anni in primissima linea sugli sviluppi dell'Industrial IoT vede nei Digital Twin un modo efficace di rispondere ai bisogni di ottimizzazione delle risorse, manifatturiere e di servizi, delle aziende. Reply e EIT Digital hanno un'attività di innovazione con lo scopo di realizzare un Digital Twin per l'anticipatory logistic (logistica predittiva) da affiancare alla tradizionale gestione di magazzino ed abilitata da smart IoT sensors.

Anche se sicuramente ai primordi, l'era dell'industria digitale e dei Digital Twin danno i primi segnali di enormi cambiamenti che sono a portata di tecnologia. Scenari dove macchine di recente installazione comunicano con altre macchine installate sul medesimo sito produttivo per cambiare la propria configurazione basandosi sull'“esperienza” accumulata da queste ultime e codificata nei relativi Digital Twin con la finalità di ridurre i fermi per manutenzione e ottimizzarne la resa complessiva, valutandone l'impatto sui propri componenti utilizzando nuovamente i Digital Twin sono già tecnicamente possibili oggi e attendono solamente di essere applicati su scala mondiale.

Tuttavia, come spesso accade nei sistemi distribuiti di prima generazione, la sicurezza sembra rappresentare un problema ancora poco affrontato. È infatti

più semplice attaccare un'entità virtuale nel cyberspace che non un asset fisico o un oggetto non "smart". Occorre quindi, nel momento in cui si propone ad un'azienda di trasferire parte dei suoi processi in un ambiente digitale ed usare i Digital Twin, garantirne la sicurezza già partendo dalla fase di progettazione. Analogamente è molto facile fare una copia di un Digital Twin, che in assenza di sistemi certificati di autenticazione potrebbe sostituire il Digital Twin primario o ereditarne importanti proprietà intellettuali di difficile protezione.

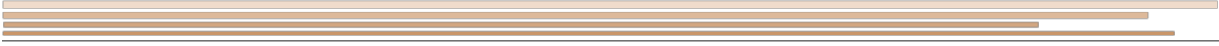
Un settore nuovo, da affrontare con gli strumenti digitali già disponibili, e integrandone ulteriori in corso di realizzazione con grandi potenzialità.

Biografia

Maurizio Griva è manager presso Concept Reply BU di Santer Reply SpA con la responsabilità di Business Development e Funded Research Projects nell'area dell'Internet of Things e Cloud/Edge computing. Ha un'esperienza pluriennale nella realizzazione di progetti di innovazione digitale precedenti in Motorola Italia, Seat PG e industrie di servizi e utility.

Email: m.griva@reply.it





Symbiotic Autonomous Systems, Digital Twins and Artificial Intelligence: Emergence and Evolution

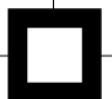
S. Mason Dambrot

Sommario

Nei prossimi anni assisteremo ad un notevole sviluppo di Sistemi Autonomi Simbiotici (SAS) e dei Digital Twins sia in termini di capacità sia in termini di applicazioni. Questa crescita sarà accompagnata, ed anche resa possibile, dalla evoluzione parallela dell'intelligenza artificiale che giungerà ad avere capacità sempre più equivalenti e quindi superiori all'intelligenza umana, ad esempio, Intelligenza Generale Artificiale (AGI) e Superintelligenza artificiale (ASI).

SAS, Digital Twin ed Intelligenza Artificiale saranno sempre più strettamente cooperanti al punto da diventare aspetti diversi di un'unica entità.

Keywords: Artificial Intelligence, Artificial General Intelligence, Artificial Superintelligence, Digital Twins, Machine Awareness, Machine Learning, Neural Networks, Symbiotic Autonomous Systems



Human cognition and creativity have manifested as our unparalleled conception, creation and usage of ever more complex and powerful tools. Our tools, in turn—along with natural environmental factors (e.g., climate, geography, and local lifeforms)—have been powerful determinants of our cultural and societal evolution. These uniquely powerful capabilities ultimately resulted from two major genetic mutations that occurred in two ancestral species in our distant past. As our toolmaking skills grew more sophisticated and functional, our control over ourselves and our surroundings increased, with our reliance on our physical corpus (perception, muscles and dexterity) and the levels of scale, mass and distance that we could interact with and modify, the key change factor being the introduction of *de novo* external energy sources modes. Mechanical calculators gave way to electronic technology, and from there to computer hardware and software.

We are now in an increasingly ubiquitous digital age that overlaps with an emerging bioelectronic period and the beginning of synthetic genetic technology that suggest a period in which we can modify our own genomes and thereby our physical bodies, including where all the above started: our brains. Beyond that lies a possible science fiction-like world in which we may be able to cure or altogether abolish disease; dramatically extend our lives; transfer our brain's content—i.e., our minds—to another neural platform, be it in a clone, computer, or synthetic neural-compatible platform.

Beyond these potential capabilities, we also may develop the ability to seamlessly share both thinking and awareness with other humans and computers, cloud databases, intelligent networks, and other seamless extensions of our conscious mind. In their most sophisticated state, these candidate Symbiotic Autonomous Systems technologies can be described as Digital Twins: digital representations of Physical Twin assets, processes, and systems. Digital Twins can represent objects and entities as varied as a turbine, a robot, internal medical nanorobotic surgeons, a vehicle of arbitrary size and complexity, a non-human mammal, a human being, an entire smart city, myriad other entities, and intangible entities such services, processes and knowledge. (At the same time, while Digital Twins mimic an object's atoms and their structural/functional relationships in bits, all bits may not be represented. However, there are upsides to this property—namely, data repositories, real-time analysis and simulation, or pattern and meaning identification.

Finally, the connections between SAS, Digital Twins and Artificial Intelligence are myriad but largely intuitive. The short version: As SAS and Digital Twins continue to develop, emerge and expand their roles, the parallel evolution of Artificial Intelligence advances with capabilities increasingly equivalent and superior to human intelligence (e.g., Artificial General Intelligence and Artificial Superintelligence), the synergy between SAS and Digital Twins with AI will become simultaneously increasingly complex yet transparent to these interdigitating components.

Our Historical Quest for Technological Intelligence

Our species is not alone in strategizing, communicating, and making and using tools. Rather, it is the degree to which we have evolved in these capacities that defines us—a result, remarkably, of two periods of neural and cranial genetic mutations. First, a replication mutation of the ARHGAP11A gene ~2–3 million years ago—that is, after our ancestors differentiated from chimpanzees, but before branching from Neanderthals—yielded the ARHGAP11B gene, which led to neocortical expansion and increased folding (Florio et al., 2015; Florio et al., 2016). This, in turn, increased our brain volume threefold (Vallender, 2008). Second, a set of genes (SRGAP2, FOXP2, and others) that may have determined changes in both neural connectivity and the shape of our skull (the latter allowing our brains, including the all-important frontal cortex, to expand), giving rise to our unique cognitive and linguistic capabilities (Benítez-Burraco et al., 2015). These modifications allowed a much greater ability to envision, convey, and create ever-increasing powerful tools.

That being said, while Artificial Intelligence (as well as its variants and related technologies) have been introduced relatively recently, precursors in the forms of myth, logic and invention are known to have existed from antiquity onwards. In ancient Greece, myths of Hephaestus and Pygmalion incorporated the idea of intelligent robots and artificial beings; Archytas of Tarentum's jump to actually constructing mechanical toys and models; and Yan Shi presented King Mu of Zhou with mechanical men. More abstract achievements in logic and mathematics—including Pythagoras' theorem, Aristotle's syllogistic logic, and Euclidian geometry—appeared between the 6th and 3rd centuries BCE, with our ability to imagine and construct ever-increasing sophisticated and diversified tools, leading at an accelerated rate to an ever greater diversity, including: moveable type printing (Bi Sheng, 1041~1048); movable type *mechanical* printing (Guttenberg, 1450); the first mechanical digital calculating machine (Pascal, 1648); a universal calculus of reasoning (Leibniz, 1673); the first punched card programmable machine (Jacquard, 1801); a programmable mechanical calculating machine (Babbage and Byron, 1832); modern propositional logic (Frege, 1879, expanded by several individuals, including Russell, Godel and Church); advanced formal logic (Russell and Whitehead, *Principia Mathematica*, 3 volumes, 1910, 1912 and 1913); universal Turing machine (Turing, proposed 1936-37); *A Logical Calculus of the Ideas Immanent in Nervous Activity* (McCulloch & Walter Pitts, 1943); *Turing Test* (Turing, 1950); *Programming a computer to play chess* (Shannon, 1950); *Three Laws of Robotics* (Asimov, 1950); the term *Artificial Intelligence* (McCarthy, Dartmouth Conference, 1956); *Logic Theorist*, the first running Artificial Intelligence program (Newell, Shaw and Simon, 1956); *ELIZA* dialogue simulation (Weizenbaum, 1965); *The Perceptron (Artificial Neural Networks)*, Rosenblatt, 1958); conceptual dependency model for natural language understanding (Schank, 1969); *SHRDLU* language understanding program with a robot arm that carried out instructions typed in English (Winograd, 1971); massively parallel Connection Machine (Hillis, 1987); *Neuromorphic electronic systems* (Mead, 1990); significant machine learning, intelligent tutoring, case-based

reasoning, multi-agent planning, scheduling, uncertain reasoning, data mining, natural language understanding and translation, vision, virtual reality, games, and other Artificial Intelligence advances (1990s); *Sociable Machines* and *KISMET* robot facially expressed emotions (Breazeal, 2000); NASA exploration Robotic rovers *Spirit* and *Opportunity* autonomously navigate the surface of Mars (NASA, 2004); *ASIMO* artificially intelligent humanoid robot walked as fast as humans and brought trays to customers (Honda, 2005); and Google Duplex Artificial Intelligence assistant schedules appointments via cellphone and produces high-quality human-sounding speech (Google, 2018)

Artificial Intelligence (AI)

In computer science, Artificial Intelligence (AI) is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals. Computer science defines AI research as the study of intelligent agents—any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals. More specifically, Kaplan and Haenlein define AI as “a system’s ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation” (Kaplan and Haenlein, 2019). Colloquially, the term Artificial Intelligence is applied when a machine mimics cognitive functions that humans associate with other human minds, such as learning and problem solving.

Moreover, augmentation is pursued through intelligence. While today we already have a number of smart machines, augmented machines will use their intelligence to augment themselves. Hence the shift from today’s local intelligence, used to be more effective in doing the activities they are supposed to do, to the next decade where machines will become smarter by using their own intelligence to tap on ambient intelligence (other machines, humans, web- and environment-distributed AI). Baxter—an industrial robot (was decommissioned and replaced by Sawyer) manufactured by Rethink Robotics (which has now closed its doors) was a first step in that direction, being able to learn by observing its co-workers. In the longer term, machines will autonomously create symbiotic relationships that realize how to best collaborate in order create a team for approaching, solving, and/or executing a task to reach a goal.

AI will continue to be applied in signal processing—a mature technology being used in a growing number of applications with increasingly complex interactions requiring signal processing to migrate from monoscale to independent multiscale and, by the second part of the next decade, simultaneous polyscale environments. AI is also progressing in bio-interfaced machines as interactions become increasingly complex (e.g., capturing and delivering electrical signals to/from thousands of probes using deep brain stimulation).

In addition to signal processing, AI is becoming increasingly ubiquitous, with applications including optical character recognition; handwriting recognition; speech recognition; face recognition; artificial creativity, computer vision, virtual reality, and image processing; photo and video manipulation; diagnosis; game theory and strategic planning; game artificial intelligence and computer game bot; natural language processing, translation and chatterbots; nonlinear control

and robotics; artificial life; automated reasoning; automation; bio-inspired computing; concept mining; data mining; knowledge representation; semantic web; email spam filtering; robotics (behavior-based robotics; cognitive; cybernetics; developmental robotics (epigenetic); evolutionary robotics); hybrid intelligent system; intelligent agents; intelligent control; and litigation.

Artificial Intelligence-related Technologies

Deep Neural Networks

Deep Neural Networks (DNN) are a layered structure of computation where each layer returns a probability that is further processed at the layer above. Probabilities are matched with the real world and change over time based on experience. Hence DNN therefore are an ideal technology for learning from experience. The tweaking of the computation may be done internally or by an external operator. Early in the next decade DNN are expected to become part of many autonomous systems, providing the capability to learn from experience, making them ever more flexible and autonomous. In the context of Symbiotic Autonomous Systems, every component, in principle, can contribute to DNN fine tuning.

Recurrent Neural Networks

Recurrent Neural Networks (RNN) are sequential structures that process and understand time evolution, and are utilized and well established in writing and speech recognition. This temporal observation is clearly relevant in Symbiotic Autonomous Systems, but while still in its early stages. RNN can be expected to become normative in the second part of the next decade.

Convolutional Neural Networks

Convolutional Neural Networks (CNN) are a class of feed-forward artificial neural networks mimicking the visual cortex in animals and are applied to image recognition. They are already part of the standard tool set for several image recognition applications. Scientists are making progress in understanding the circuitry of animals' brain, like the brain of a fly, and are investigating the effectiveness of replicating their capabilities in artificial neural networks. With a relative limited number of neurons and very little power requirements, a fly can orient itself in a 3D space whereas our artefacts require a massive amount of processing.

Machine Learning

Machine Learning is a mature technology, in the sense that it is already widely used today. Nevertheless, we can expect significant growth in its capability, due to neuromorphic chips (i.e., analog processors based on biological neurons) with increased capability and the extension of the data the it uses to increase learning-areas where industry research will be leading. By 2030, machine learning is expected to be a standard component in most systems. (Note that by the end of the next decade, neuromorphic chips supporting AI toolkits can be expected). Around 2040, a new generation of neuromorphic chips, at very low power (getting closer to bio-neural circuits in terms of power requirement) may become available as a platform to support Artificial General Intelligence (see the AGI section below). These chips will be de facto implemented in their

architecture like today is done by deep/recurrent/convoluted neural networks in software running on normal chips. In turns this will make machine awareness possible and affordable.

Moreover, there will be an evolution of this functional area from today's *independence* in which the machine operates independently of the host biosystem (e.g., a pacemaker that sends impulses to the heart without being aware of the body's general situation) to *responsiveness*, where the machine senses the status of the bioentity and adapts its actions as needed) to a *continuous interaction*, and finally to a *symbiotic status* where machine and bioentity influence each other towards a common goal.

Machine Awareness

In order to get smarter, machines need to become more and more aware of their context, goals, and abilities. By far the basic enabling technologies are the capability to process the signals received through sensors from the environment (including their active observation of the environment) and the intelligence to make sense out of them. Accordingly, three phases can be identified: task awareness, goal awareness, and self-awareness.

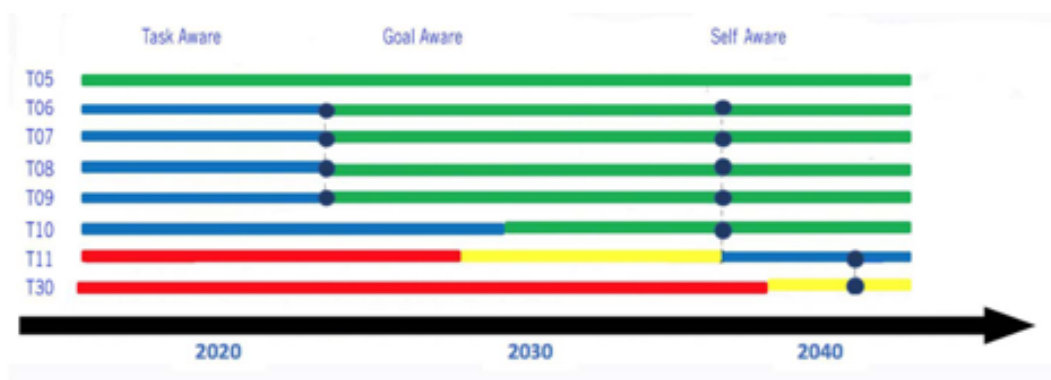


Figure 1
Timeline of context-aware machine-related technologies

Legend:

- T05: Signal processing
- T06: Artificial Intelligence
- T07: Deep Neural Networks
- T08: Recurrent Neural Networks
- T09: Convoluted Neural Networks
- T10: Machine Learning
- T11: Artificial General Intelligence
- T30: Sentient Machines

Context-Aware Machines

As machines can harvest and process more data from their environment to create a model of the environment and to perceive their role and interaction with

the environment, they are shifting from being passive to becoming active towards an understanding of the environment. There are a few areas that are already seeing this evolution with self-driving cars at the forefront (it is likely that in the military area there is faster evolution, but progress is not disclosed). The context-awareness has already reached the mass market in products like robotic vacuum cleaners, but it is focused on very specific environment niches. A more generalized context-awareness will take a few more decades to become the norm.

Machine Sentience

It is generally accepted that in order to be sentient an entity should be able to sense, perceive, think, feel and experience subjectivity. Unsurprisingly, then, if there is not a consensus on machine awareness metrics, it is not unexpected that there is even less so regarding sentient machines—that is, machines that have a self and recognize themselves as a “living” entity with a goal and a sense of fulfilment. That said, there are technologies that clearly cover the sense and perception aspect; support processing that can be described as thinking; and machines that can be programmed with a sense of identity (subjectivity). The difficult part, however, is related to feelings: While a machine can be programmed to respond to rewards, as well as having a sense of an undesirable situation, it is currently unknowable if these machine states translate into feelings or emotions. These are complex states that require the participation of several biological structures and functions in the mammalian brain, including the frontal lobe, the olfactory bulb, the thalamus and hypothalamus, the hippocampus, and the amygdala—all parts of the limbic system. While a hybrid robot might *simulate* such processes, it would remain to be seen whether that would lead to *experience* the feelings necessary to correlate with the attending circumstances.

Affective Computing

An endeavor complementing and relating to AI/AGI, affective computing is the development of technology capable of recognizing, interpreting, processing, and simulating (and eventually actually possessing) human or human-analogous affect, the goal being to provide “computing that relates to, arises from, or influences emotions” (Picard, 1995), and realizing that while emotional arousal data has been collected for over a century, “technology can now measure, communicate, adapt to, be adapted by, and transform emotion. (Picard, 2010). The MIT Affective Computing Group is applying these capabilities to “help people gather, communicate, and express emotional information and to better manage and understand the ways emotion impacts health, social interaction, learning, memory, and behavior” (MIT Affective Computing Group).

Essentially, affective computing seeks to give computers the capability to take into account emotion related to human cognition and perception, a key result being Digital Twins with the capacity to more effectively support humans, as well as they themselves being able to make better decisions. Current investigations into applying affective computing in the enhancement of applications include computer-assisted learning, perceptual information retrieval, and human health and interaction.

Artificial General Intelligence (AGI)

While Artificial Intelligence can be said to have a rudimentary, narrow awareness of basic environmental activity or changes (for example, signal frequencies and document edits), it will be Artificial General Intelligence with human-like characteristics of self-awareness and conscious reflection that will emerge. In this sense, AGI is rapidly becoming a mature, industrial grade technology expected to become a common tool in industry by the middle of the next decade—and as such, will likely be a subsystem that is invoked as needed within AGI environments.

Artificial General Intelligence refers to a machine intelligence that can successfully perform any human intellectual task. AGI is a primary goal of some AI research and a common topic in future studies and science fiction. (Note that some researchers refer to Artificial General Intelligence as *strong* AI, others assign the term AGI to machines capable of experiencing consciousness.

At the same time, AGI is not a requirement for achieving machine awareness—and it is not given that AGI will become possible by 2050 (although some researchers are convinced that this will take place within the next 20 years). In fact, there is no consensus on the achievability of AGI, its unique definition, or metrics for determining when AGI has become a reality. Therefore, in the absence of an unexpected breakthrough, AGI will primarily be a matter of academic research for the next decade—and while a few companies, such as Google, are involved, general industry is not expected to become active in AGI before the 2040 timeframe. That said, acceleration may derive from results of ongoing brain projects—for example, Human Brain6, Brain Initiative7, and Human Connectome Project—and as discussed above, AGI is not required to support basic machine awareness, with AGI demonstrating self-awareness.

Even with the above considerations, AI has progressed in the last decade, due largely to vast computation capabilities and access to Big Data in specific areas (e.g., speech recognition and understanding). However, creating AGI has proven elusive to the point that some experts are not optimistic on achieving it in the coming decades—but others assert that it will be a reality by 2030. In reality, it is becoming difficult to place a boundary between AI and AGI (for example, will AGI include emotions and self-awareness?). AI is bound to extend its fields of application to the point at which it might be practically indistinguishable from AGI.

A development on the forefront of AGI research, startup Artificial General Intelligence Inc. recently published a paper that articulates the fundamental theory of consciousness used in the Independent Core Observer Model (ICOM) research program (Kelley and Twyman, 2018) and the consciousness measures as applied to ICOM systems and their uses in context including defining the basic assumptions for the ICOM Theory of Consciousness (ICOMTC) and associated related consciousness theories (CTM, IIT, GWT, etc.) upon which ICOMTC is built. The paper defines the contextual experience of ICOM based systems in terms of a given instances subjective experience as objectively measured and the qualitative measure of Qualia in ICOM based systems, stating that ICOM is a cognitive architecture based on the ICOM Theory of Consciousness for independent self-aware intelligence that is able to experience

things emotionally with an internal subjective experience, making decision purely based on how it “feels” about a thought (Kelley and Waser, 2018).

The paper notes that the Independent Core Observer Model Theory of Consciousness is partially built on the Computational Theory of Mind, where an AGI research core issue is the absence of objective consciousness: The related measurements are ambiguous due to the lack of agreed-upon objective measures of consciousness. Therefore, to continue serious work in the field we need to be able to measure consciousness in a consistent way that is not presupposing different theories of the nature of consciousness and further not dependent on various ways of measuring biological systems. Rather, the focus must be on the elements of a conscious mind in the abstract—and with the more nebulous Computational Theory of Mind, research into the human brain does show some underlying evidence to the same.

The authors made several assumptions to provide an experimental reference point when designing AGI cognitive architecture:

- *Qualia is the subjective experience that can be measured external to the system if the mind in question is operating under known parameters.*
- *Humans are not able to make logical decisions.* Looking at the neuroscience behind decisions the researchers have already proven that humans make decisions based on how they feel and not based on logic.
- *Subjective experience can be measured and understood.* As a matter of principle in this paper, the traditional view that the subjective nature of experience is purely subjective is rejected. Even scientists in the field frequently consider consciousness the realm of ontology and therefore philosophy and religion. The researchers’ assumption is that this is false.
- *Consciousness can be measured.* Despite this enormous commitment to the study of consciousness on the part of cognitive scientists covering philosophical, psychological, neuroscientific and modelling approaches, until now no stable models or strategies for the adequate study of consciousness have emerged, the exception being the ICOMTC. The researchers also believe that they can measure consciousness regarding task accuracy and awareness as a function of stimulus intensity that applies to brain neurochemistry as much as the subjective experience from the point-of-view of systems such as ICOM.
- *The researchers have a concrete definition of “subjective” as a concept.* “Subjective” is then defined as the relative experience of a conscious point of view that can only be measured objectively only from outside the system, where the system in question experiences things “subjectively” as they relate to that systems internal emotional context.
- *Consciousness is a system that exhibits the degrees or elements of the Porter method for measuring consciousness regarding its internal subjective experience.* While the dictionary might define consciousness subjectively in terms of being awake or aware of one's surroundings, this is a subjective definition—but the researchers need an “objective” definition

to measure. Therefore, the latter is assumed for the context of ICOM research and the ICOMTOC.

Artificial Superintelligence (ASI)

Artificial Superintelligence is a hypothetical agent that possesses intelligence far surpassing that of the brightest and most gifted human minds. (University of Oxford philosopher Nick Bostrom defines superintelligence as "any intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest" (Bostrom, 2014). Artificial Superintelligence (ASI) is artificial intelligence that surpasses the brightest human minds in all areas., thereby surpassing human intelligence-equivalent Artificial General Intelligence. Moreover, several researchers observe that computers today are better than humans in several areas (e.g., calculus) and are getting increasingly better in a range of fields. This implies that once AI reaches the AGI, it may become an ASI—thereby bypassing the AGI transition. In this scenario, machines will not achieve human intelligence equivalence, segueing directly from human intelligence-inferior to human intelligence-superior. Note, however, that ASI does *not* imply Artificial Consciousness.

Technological researchers disagree about how likely present-day human intelligence is to be surpassed. Some argue that advances in "narrow" artificial intelligence (AI) will probably result in general reasoning systems that lack human cognitive limitations. Others believe that we will evolve or directly modify our biology to achieve radically greater intelligence. A number of futures studies scenarios combine elements from both of these possibilities, suggesting that humans are likely to interface with computers, or upload their minds to computers, in a way that enables substantial intelligence amplification.

Superintelligence may also refer to a property of problem-solving systems (e.g., superintelligent language translators or engineering assistants) whether or not these high-level intellectual competencies are embodied in agents that act in the world.

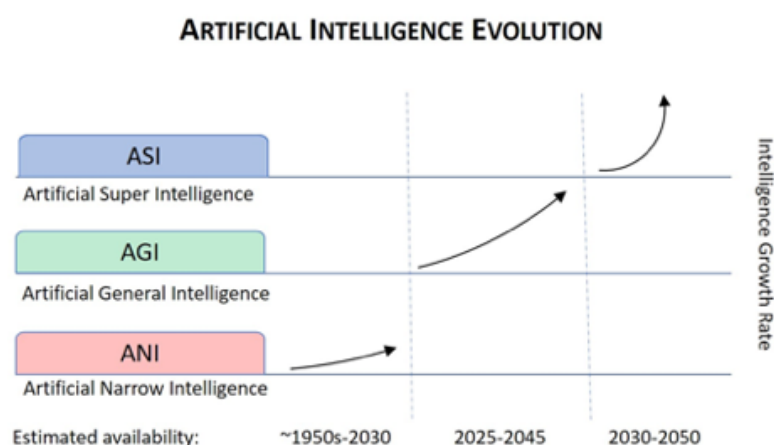


Figure 2
Artificial Intelligence Evolution

Some researchers believe that Artificial Superintelligence will likely follow shortly after the development of Artificial General Intelligence. The first generally intelligent machines are likely to immediately hold an enormous advantage in at least some forms of mental capability, including the capacity of perfect recall, a vastly superior knowledge base, and the ability to multitask in ways not possible to biological entities. This may give them the opportunity to—either as a single being or as a new species—become much more powerful than humans, and to displace them.

In a survey of the 100 most cited authors in AI (as of May 2013 (Müller et al., 2016), according to Microsoft academic search), the median year by which respondents expected machines "that can carry out most human professions at least as well as a typical human" (assuming no global catastrophe occurs) with 10% confidence is 2024 (mean 2034, st. dev. 33 years), with 50% confidence is 2050 (mean 2072, st. dev. 110 years), and with 90% confidence is 2070 (mean 2168, st. dev. 342 years). These estimates exclude the 1.2% of respondents who said no year would ever reach 10% confidence, the 4.1% who said 'never' for 50% confidence, and the 16.5% who said 'never' for 90% confidence. Respondents assigned a median 50% probability to the possibility that machine superintelligence will be invented within 30 years of the invention of approximately human-level machine intelligence.

Mediated Artificial Superintelligence (mASI)

Mediated Artificial Super Intelligence is an incremental in-process AI > AGI > ASI evolution goal. Demonstrated in the lab, and usable now in environments from research to business, mASI provides ASI superhuman level cognition without ethical or safety concerns—and markedly reduces training time (Jangra et al., 2013). mASI requires human support at all times to mediate the process to the degree that its thinking and operations do not function without human involvement. The mASI cognitive architecture the researchers are using is based on the ICOM Theory of Consciousness (Kelley and Twyman, 2018), which itself is based on Global Workspace Theory (Baars), the Computational Theory of Mind (Rescorla, 2016), and Integrated Information Theory (Tononi, 2014)—and at some level is demonstrably conscious (Yampolskiy, 2019) (Kelley, peer review pending).

Envisioning Far-Future Symbiotic Autonomous Systems Artificial Intelligence

Looking further into the AI future within an SAS context, and assuming that the increasingly human-like intelligence demonstrated in AGI and ASI will continue and likely accelerate, the theoretical systems proposed in Table 1 (Dambrot, 2019) may be seen as a feasible vision of far-future Artificial Superintelligence variants. Perhaps the most powerful concept is the Quantum Computing Artificial Superintelligence (qASI) variant, given the hypothetical properties of such a system (quantum entanglement-based communications, superposition-based parallel processing, superluminal instantaneous processing speed, and infinite number of threads). Such a massively distributed real-time system could enable space-and time-agnostic metahuman intelligence without the characterizing current systems.

Table 1

Proposed Mediated ASI	Key Properties	Benefits
<i>Distributed Artificial Superintelligence (daisy)</i>	Networked independent ASI nodes form a collective mind	• Multifocus distributed superintelligence system far beyond superhuman AGI • Distinctly nonhuman-like superintelligence • Reduced footprint
<i>Genetic Computing Artificial Superintelligence (gASI)</i>	Genetic computing-based sASI	• Increased mediation speed • CRISPR-based modification
<i>Quantum Computing Artificial Superintelligence (qASI)</i>	Quantum entanglement-based sASI	• Quantum entanglement-based communications • Superposition-based parallel processing • Superluminal instantaneous processing speed • Infinite number of threads

Conclusion

The accelerating evolution in Symbiotic Autonomous systems, Digital Twins and increasingly human-equivalent and superior generations of Artificial Intelligence—and the synergy between these three domains—present an expanding integration of biology, technology and biotechnology in which separations between these now distinct classes of entities will gradually disappear. As these trends continue, expand and accelerate, it will be just a matter of time for this state of affairs to become a forgotten memory.

References

- Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Progress in Brain Research*. Elsevier, Volume 150, pp. 45-53. <http://www.sciencedirect.com/science/article/pii/S0079612305500049>
- Benítez-Burraco, A. and Boeckx, C. (2015). Possible functional links among brain- and skull-related genes selected in modern humans. *Frontiers in Psychology* 6:794. <https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00794/full>
- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press. https://www.amazon.com/dp/B00LOOCGB2/ref=dp-kindle-redirect?_encoding=UTF8&btkr=1
- Dambrot, S. M. (2019). Proposed Far-Future Artificial Superintelligence Variants. Independent research.

- Florio, M., Albert, M., Taverna, E., Namba, T., Nüsslein, I., et al. (2015). Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science* 347(6229):1465-70. <http://science.sciencemag.org/content/347/6229/1465>
- Florio, M., Namba, T., Pääbo, S., Hiller, M., and Huttner, W. B. (2016). A single splice site mutation in human-specific ARHGAP11B causes basal progenitor amplification. *Science Advances* 2(12): e1601941. <http://advances.sciencemag.org/content/2/12/e1601941>
- Jangra, A., Awasthi, A., and Bhatia, V. (2013). A Study on Swarm Artificial Intelligence. *IJARCSSE v3 #8*. https://www.researchgate.net/publication/308719281_A_Study_on_Swarm_Artificial_Intelligence/
- Kaplan, A. and Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*. Vol. 62, Issue 1, pp. 15-25. <http://www.sciencedirect.com/science/article/pii/S0007681318301393>
- Kelley, D. J. Architectural Overview of a "Mediated" Artificial Super Intelligent Systems based on the Independent Core Observer Model Cognitive Architecture. Submitted 1 October 2018 to *Informatica Journal* (peer review pending).
- Kelley, D. J. and Waser, M. R. (2018). Human-like Emotional Responses in a Simplified Independent Core Observer Model System. Volume 123, pp. 221-227. <http://www.sciencedirect.com/science/article/pii/S1877050918300358>
- Kelley, D. J. and Twyman, M. (2018). Independent Core Observer Model (ICOM) Theory of Consciousness as Implemented in the ICOM Cognitive Architecture and the Associated Consciousness Measures. *AAAI Spring Symposia 2019*.
- Müller, V. C. and Bostrom, N. (2016). Future Progress in Artificial Intelligence: A Survey of Expert Opinion. In Müller, Vincent C., *Fundamental Issues of Artificial Intelligence*. Springer. pp. 553–571. <http://www.philpapers.org/archive/MLLFPI>
- Mead, Carver (1990). Neuromorphic electronic systems. *Proceedings of the IEEE*. 78 (10): 1629–1636. <http://ieeexplore.ieee.org/abstract/document/58356/>
- MIT Affective Computing Group. <https://www.media.mit.edu/groups/affective-computing/overview/>
- Monroe, D. (2014). Neuromorphic computing gets ready for the (really) big time. *Communications of the ACM*. **57**(6): 13–15. doi:10.1145/2601069
- Picard, R. W. (1995). Affective Computing. MIT Media Laboratory Perceptual Computing Section Technical Report No. 321. [https://www.pervasive.jku.at/Teaching/_2009SS/SeminarausPervasiveComputing/Begleitmaterial/Related%20Work%20\(Readings\)/1995_Affective%20computing_Picard.pdf](https://www.pervasive.jku.at/Teaching/_2009SS/SeminarausPervasiveComputing/Begleitmaterial/Related%20Work%20(Readings)/1995_Affective%20computing_Picard.pdf)
- Picard, R. W. (2010). Affective Computing: From Laughter to IEEE. *IEEE Transactions on Affective Computing*, Vol. 1, No. 1. <https://affect.media.mit.edu/pdfs/10.Picard-TAC.pdf>
- Rescorla, M. (2016). Computational Theory of Mind. Stanford University. <http://plato.stanford.edu/entries/computational-mind/>

Symbiotic Autonomous Systems White Paper I <https://symbiotic-autonomous-systems.ieee.org/images/files/pdf/sas-white-paper-final-nov12-2017.pdf>

Symbiotic Autonomous Systems White Paper II <https://symbiotic-autonomous-systems.ieee.org/images/files/pdf/SAS-WP-II-2018-Finalv3.2.pdf>

The Human Brain Project SP 9: Neuromorphic Computing Platform <https://www.youtube.com/watch?v=6Roiz90mGfw>

Tononi, G., Albantakis, L., and Masafumi, O. (2014). From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. Computational Biology. <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003588>

Vallender, E.J., Nitzan, M-B., and Lahn, B.T. (2008). Genetic basis of human brain evolution. *Trends in Neurosciences* 31(12):637-44. <https://doi.org/10.1016/j.tins.2008.08.010>

Yampolskiy, R. V. (2019). Artificial Intelligence Safety and Security. Chapman and Hall/CRC Artificial Intelligence and Robotics Series, London/New York. ISBN: 978-0815369820. https://www.amazon.com/Artificial-Intelligence-Security-Chapman-Robotics-ebook-dp-B07G19SGK7/dp/B07G19SGK7/ref=mt_kindle?_encoding=UTF8&me=&qid=

Zhao, W. S., Agnus, G., Derycke, V., Filoramo, A., Bourgoin, J. -P., and Gamrat, C. (2010). Nanotube devices based crossbar architecture: Toward neuromorphic computing. *Nanotechnology*. 21 (17): 175202. <https://iopscience.iop.org/article/10.1088/0957-4484/21/17/175202/meta>

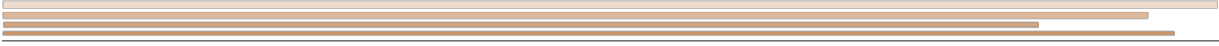
Biografia

S. Mason Dambrot is a Transdisciplinary Researcher | Theorist, Research Fellow at Brain Machine Interface Consortium, Research Fellow at Artificial General Intelligence Society, and Board Member at Artificial General Intelligence Inc. His IEEE memberships include Symbiotic Autonomous Systems Initiative; Engineering in Medicine and Biology Society; Systems, Man and Cybernetics Society; Standards Association; and Brain Community. He serves on the IEEE Symbiotic Autonomous Systems Steering Committee and IEEE SmartAg Executive Committee.



His recent publications include *ReGene: Blockchain backup of genome data and restoration of pre-engineered expressed phenotype* (IEEE Xplore, In Press); *Enplants: Genomically engineered neural tissue with neuroprosthetic and communications functionality* (IEEE Xplore, 2018); *Exocortical Cognition: Heads in the Cloud—A transdisciplinary framework for augmenting human high-level cognitive processes* (IEEE Xplore, 2017); and *Neuroprosthetics: Past, Present, and Future*, in *Brain Computer Interfaces Handbook: Technological and Theoretical Advances* (Taylor & Francis Group, 2018).

Email: smdambrot@ieee.org



Sistemi Autonomici - Aspetti di comunicazione

Massimo Donelli

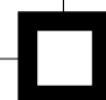
Sommario

Lo sviluppo di sensori miniaturizzati, in grado di rilevare e trasmettere informazioni tramite un canale wireless, ha reso possibile la crescita dell'Internet delle cose IoT. Grazie all'IoT, si stanno creando le condizioni ottimali per l'espansione di reti estremamente pervasive e complesse, che possono fornire differenti tipologie di servizi e offrire soluzioni a complicati problemi tecnici. Tuttavia, questa rapidissima evoluzione richiede lo sviluppo di sistemi capaci di adattarsi ad essere operativi in differenti scenari caratterizzati da contesti temporali dinamici; nello stesso tempo, questi sistemi devono avere la capacità di autoadattarsi in termini di consumo energetico, connettività, oltre ad essere eventualmente in grado di autoripararsi. La sfida è implementare un sistema autonomico. In tale contesto la connettività dei vari elementi di tale rete distribuita è di fondamentale importanza. In questo articolo si cercherà di analizzare le problematiche dei collegamenti wireless in termini di sistemi autonomici.

Keywords: Wireless systems, smart antenna, optimization techniques, unsupervised control methods, digital transformations

Introduzione

Il termine sistema autonomico trae le sue origini dalla terminologia medica, in particolare dalla fisiopatologia. Un classico esempio è il sistema nervoso autonomico, che è un sistema capace di reagire in maniera reversibile ad un



impulso nervoso o ad uno stimolo proveniente dall'ambiente esterno. Il sistema nervoso autonomico si occupa del controllo del muscolo cardiaco, della muscolatura liscia e delle ghiandole esocrine. Esercita il ruolo principale del nostro sistema vegetativo, regolando le funzioni vitali interne (temperatura, battito cardiaco) in base agli stimoli provenienti dall'esterno. La parte di sistema nervoso denominata ortosimpatico presiede alle reazioni di lotta e fuga, mentre il parasimpatico governa il riposo e la digestione. Tutte queste funzioni di fondamentale importanza avvengono in maniera automatica e senza che l'individuo, ne abbia coscienza. E' pertanto un sistema non supervisionato. Nel lontano 2001, il termine autonomico è stato preso in prestito da IBM che ha introdotto la cosiddetta computazione autonomica "Autonomic Computation", con lo scopo di creare un sistema distribuito in grado di far fronte alla crescente complessità ed eterogeneità dei sistemi informatici. Grazie ai recenti avanzamenti tecnologici, tali sistemi distribuiti hanno raggiunto un'elevata pervasività e possono coprire aree molto vaste, come un'intera città. In tale contesto, gli aspetti di comunicazione tra gli elementi costituenti il sistema autonomico giocano un ruolo di fondamentale importanza. Un sistema autonomico è una rete costituita da vari elementi, detti unità fondamentali o nodi, che comunicano e interagiscono tra di loro, mutando il proprio comportamento e le proprie relazioni con le altre unità, seguendo regole codificate all'interno di ogni unità. In generale, ogni singola unità di un sistema autonomico esegue due cicli di controllo, uno interno e uno esterno. Il controllo interno serve per la gestione interna del componente (ottimizzazione delle risorse energetiche, identificazione e riconfigurazione in caso di anomalie o guasti, ottimizzazione dei parametri di comunicazione); il secondo ciclo effettua l'interazione con l'ambiente esterno, garantendo l'adattamento allo scenario operativo. I due cicli di controllo sono schematizzati in Fig. 1. Ogni componente di un sistema autonomico è quindi un'entità in grado di raccogliere gli stimoli provenienti dall'ambiente esterno o forniti da altri componenti e di auto-adattare il proprio comportamento tenendo conto delle variazioni del proprio stato interno e delle interazioni con altri componenti. Il componente elementare applica logiche di autotutela, per ottimizzare le proprie risorse, auto-configurarsi, al fine di aumentare il proprio tempo di vita. Tali logiche permettono alle componenti del sistema autonomico di auto-adattarsi alle variazioni, al fine di evitare o risolvere situazioni di malfunzionamento, errore, proteggersi da possibili eventi critici e ottimizzare le proprie prestazioni. Un efficiente sistema di comunicazione wireless è di fondamentale importanza, specie per l'esecuzione del ciclo di controllo esterno, che deve garantire in ogni condizione il corretto cambio d'informazione al fine di riorganizzare al meglio il sistema autonomico. In questo articolo ci focalizzeremo sulle tecniche di comunicazione wireless e sui sistemi radianti in grado di autogestirsi ed auto-configurarsi, al fine di ottimizzare la comunicazione tra i vari elementi della rete distribuita. In particolare, verranno analizzate tecniche che si ispirano agli sciame intelligenti di insetti quali api, formiche, lucciole, e alla stessa evoluzione darwiniana. Queste tecniche hanno ispirato la realizzazione dei sistemi autonomici.

**Figura 1**

Sistemi autonomici schematizzazione dei due cicli di controllo interno ed esterno.

Implementazione di un sistema di comunicazione con capacità autonome basate su comportamenti naturali

Come accennato nell'introduzione, in un sistema autonomico le unità elementari, dette nodi, devono essere in grado di poter comunicare tra di loro, scambiare dati e informazioni, al fine di organizzarsi e assolvere specifici compiti. In tale contesto, la comunicazione tra i vari nodi è di fondamentale importanza e, considerando che i sistemi autonomici possono essere distribuiti su un'area molto vasta, come per esempio una città o una vasta area rurale, la comunicazione deve essere effettuata per mezzo di un canale wireless. Al fine di creare una rete di comunicazione efficiente diverse, topologie di reti possono essere considerate: in particolare, le topologie più comuni sono illustrate in Fig. 2; le topologie a bus e completamente connesse sarebbero le più adatte per l'implementazione di un canale di comunicazione adatto ad un sistema autonomico. Tuttavia, le topologie illustrate in Fig. 2 sono di difficile, se non impossibile, implementazione e certamente poco adatte ad uno scenario tempo-variante. Prendiamo come esempio una rete di sensori atti al rilevamento dell'inquinamento atmosferico in un'area urbana densamente popolata e distribuita su parecchi chilometri quadrati. Tale scenario risulta essere un classico esempio di ambiente con caratteristiche fortemente variabili nel tempo, basti pensare al traffico, al passaggio di merci o persone dovute alle comuni

attività giornaliere; in particolare il passaggio di un autoveicolo può interrompere la comunicazione tra due nodi. L'organizzazione delle comunicazioni tra i vari elementi della rete necessita di una gestione complicata, in grado di garantire in ogni condizione la comunicazione tra i vari elementi della rete e il trasferimento dell'informazione verso il sistema di immagazzinamento ed elaborazione. La gestione delle comunicazioni wireless e dei collegamenti tra gli elementi base del sistema può essere efficacemente implementata creando un sistema autonomico ispirato agli algoritmi evolutivi basati sull'evoluzione darwiniana o sugli sciame intelligenti. La Fig. 3 illustra gli algoritmi evolutivi più utilizzati e diffusi. Tali algoritmi agiscono come un sistema autonomico. Gli algoritmi evolutivi si suddividono in due classi: quelli competitivi e quelli collaborativi. Gli algoritmi competitivi mimano l'evoluzione darwiniana, simulano una popolazione di individui che si evolve nel tempo seguendo le regole dell'evoluzione: due individui della popolazione vengono selezionati (denominati genitori) si accoppiano in modo da scambiarsi il codice genetico e subiscono mutazioni casuali; i nuovi individui generati, detti figli, se hanno caratteristiche che si adattano meglio alle condizioni ambientali, prendono il posto dei genitori nella popolazione. La popolazione si auto organizza seguendo le semplici regole dell'evoluzione darwiniana, basata su tre operazioni fondamentali: la selezione, l'accoppiamento con scambio di codice genetico detta cross-over e la mutazione. Gli individui della popolazione evolvono in maniera non supervisionata. L'evoluzione darwiniana è un processo competitivo: gli individui della popolazione, come in natura, competono per trasmettere il proprio codice genetico. Questa competizione porta ad un fenomeno detto stagnazione: nella stagnazione il codice genetico della popolazione risulta essere piuttosto uniforme, questo conduce ad avere nuovi individui figli con lo stesso codice genetico dei genitori, pertanto non è più possibile nessun miglioramento evolutivo. L'altra classe di algoritmi evolutivi è quella basata sugli sciame intelligenti. Gli sciame intelligenti mimano comportamenti di esseri viventi caratterizzati da un comportamento sociale, quali sciame di insetti (api, lucciole, formiche) o banchi di pesci, stormi di uccelli e pipistrelli. Gli individui dello sciame agiscono indipendentemente, seguendo semplici regole comportamentali, e collaborano tra di loro scambiandosi informazioni, in modo tale da perseguire uno scopo comune. Nelle api, formiche e negli stormi di uccelli lo scopo comune è la ricerca del cibo nel caso delle lucciole, l'accoppiamento. Grazie alle semplici regole di comportamento seguite dagli individui dello sciame e al periodico scambio delle informazioni, lo sciame è in grado di auto-organizzarsi e rappresenta di per sé un sistema autonomico. Vediamo come tali algoritmi possono essere utilizzati per la gestione di un sistema autonomico; in particolare gli algoritmi evolutivi possono essere di grande utilità per la gestione della comunicazione tra le unità elementari di un sistema o una rete autonoma. Le unità elementari vengono codificate matematicamente come individui della popolazione (per gli algoritmi competitivi) e membri dello sciame, nel caso degli algoritmi collaborativi. L'obiettivo da raggiungere viene codificato mediante una funzione matematica denominata funzionale di costo. Cerchiamo di spiegare con un esempio applicativo come tali sistemi possano essere efficacemente utilizzati per la gestione e

l'organizzazione di un sistema autonomico. Si consideri l'esempio precedente, consistente in una rete di N sensori per il rilevamento dell'inquinamento atmosferico distribuiti su una vasta area urbana densamente popolata. L'obiettivo di ogni sensore è quello di inviare i dati relativi ad un'area monitorata verso un database centrale, in modo che possano essere immagazzinati e processati. I singoli sensori rappresentano le unità elementari del sistema autonomico; a causa della vastità (estesa per parecchi Km) e della pervasività della rete dei sensori non possono avere un collegamento wireless diretto gli uni con gli altri. Considerando che i dispositivi di comunicazione wireless montati sui sensori hanno una portata limitata a qualche decina di metri bisogna organizzare la rete di sensori in modo tale che sia in grado di auto-organizzarsi, al fine di inviare l'informazione a destinazione. Tale tipologia di problema non è di facile soluzione e, se effettuato in modo supervisionato richiederebbe un notevole impegno in termini di tempo e di risorse utilizzate. La gestione di tale sistema può essere effettuata ispirandoci al mondo animale e agli algoritmi evolutivi, in particolare alle colonie di formiche e all'algoritmo "Ant Colony". Vediamo come funziona l'algoritmo "Ant Colony", ispirato al comportamento delle colonie di formiche legionarie. Le colonie di formiche sono in grado di auto-organizzarsi, utilizzando semplici segnali chimici. In particolare, quando una formica che si muove autonomamente lascia il nido e trova una locazione ricca di cibo, rilascia una sostanza chimica chiamata ferormone che attira le altre compagne che a loro volta rilasciano feromone, evidenziando un percorso ben stabilito. Aree con poca disponibilità di cibo vengono poco frequentate, il feromone evapora, cancellando la traccia. Se si osserva un nido di formiche si può notare che i percorsi delle formiche risultano piuttosto caotici.

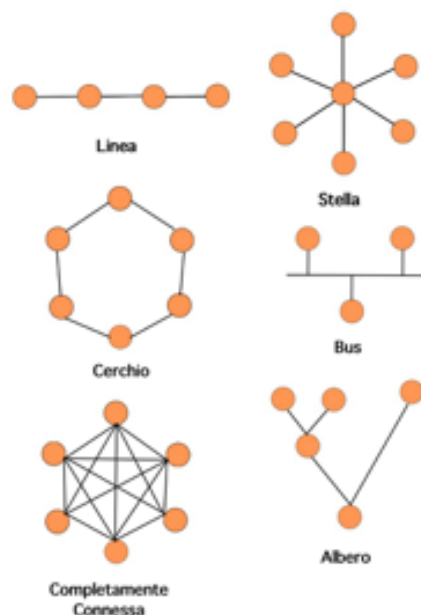


Figura 2
Topologie di reti più utilizzate.

Dopo un po' di tempo, si osserva che le formiche si organizzano per percorsi ben definiti e poi le formiche marciano allineate come soldati (da qui il nome formiche legionarie), utilizzando percorso che va dal nido fino ad una certa locazione. Vediamo come questo comportamento autonomico può essere applicato alla gestione della rete di sensori posizionata nell'area urbana densamente popolata. Il nido rappresenta la stazione dove devono essere immagazzinati e processati i dati, e le unità fondamentali della rete rappresentano la singola formica. Supponiamo di dover trasferire un insieme di dati da un sensore che chiameremo M al nido N. Il trasmettitore montato a bordo del singolo sensore (formica) ha una portata molto limitata e può comunicare solo con le formiche nelle immediate vicinanze. Pertanto è necessario individuare un percorso preferenziale tra i vari nodi, così da trasferire l'informazione in modo rapido e sicuro. Ogni formica deve ricevere feedback dall'ambiente esterno, in particolare conoscere quante formiche si trovano nelle sue vicinanze e la bontà del canale wireless che può essere facilmente misurata tramite un parametro chiamato rapporto segnale rumore. Naturalmente, data la natura variabile dello scenario, il percorso individuato dalle formiche in un certo periodo di tempo può diventare inefficiente o inutilizzabile; a questo punto la rete si riorganizzerà in modo da trovare un percorso alternativo che garantisca il corretto invio dell'informazione. Le singole formiche hanno una capacità computazionale molto limitata pertanto non possono gestire algoritmi molto complessi o immagazzinare troppe informazioni. Grazie all'implementazione ispirata al comportamento delle formiche l'unica elaborazione che il nodo può effettuare è la ricerca di un nodo adiacente con un buon collegamento radio. La rete di misura distribuita che abbiamo appena descritto rappresenta un sistema autonomico in grado di organizzarsi e trasmettere correttamente l'informazione da un nodo (formica) alla stazione base (nido), ed è stata completamente ispirato al comportamento sociale delle formiche. La Fig. 4 schematizza il sistema autonomico appena descritto. Per maggiori approfondimenti sugli algoritmi evolutivi ispirati ad esseri viventi dal comportamento sociale e alle loro applicazioni a problemi complessi di ingegneria, si prega di fare riferimento alla sezione dedicata ai riferimenti bibliografici.

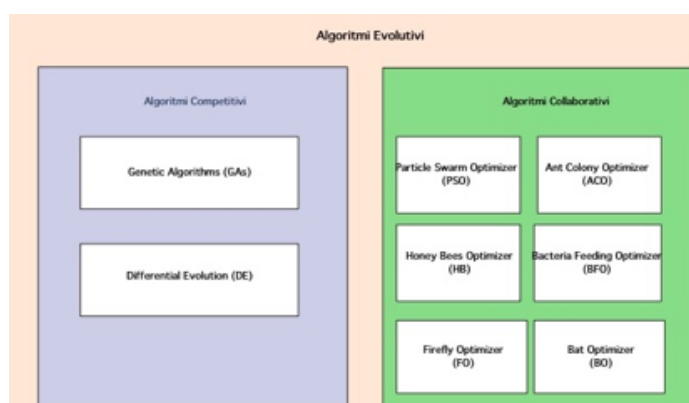


Figura 3
Algoritmi evolutivi più diffusi.

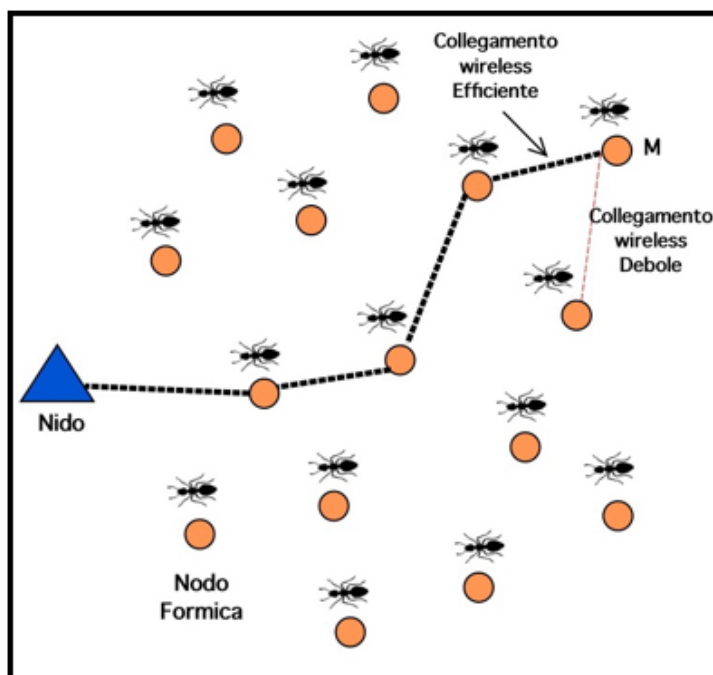


Figura 4

Esempio di rete di monitoraggio distribuita autonoma organizzata simulando il comportamento sociale delle formiche.

Altri esempi di Sistemi di comunicazione con capacità autonome: le antenne intelligenti

L'evoluzione dei componenti elettronici ha reso disponibili dispositivi compatti in grado di operare ad elevate frequenze, ed ha reso possibile la realizzazione di sistemi radianti intelligenti, ovvero antenne in grado di adattare in tempo reale le proprie caratteristiche radiative, al fine di massimizzare il raggio operativo, di ottenere un buon rapporto segnale rumore e garantire un canale wireless efficiente in qualsiasi condizione. Le antenne intelligenti sono solitamente realizzate mediante antenne a cortina. Le antenne a cortina sono costituite da diversi elementi elementari dello stesso tipo e tutti orientati nella stessa direzione. Ogni elemento è alimentato indipendentemente tramite attenuatori e sfasatori programmabili, al fine di modificare le proprie caratteristiche radiative. Lo schema di un'antenna intelligente è mostrato in Fig. 5. Agendo sulle fasi e sulle ampiezze dei singoli elementi, è possibile dirigere il lobo principale dell'antenna lungo una direzione ben definita, al fine di ottenere un collegamento radio in linea di vista ottimale. Le antenne intelligenti sono in grado di inserire dei nulli nel diagramma di radiazione, con lo scopo di eliminare eventuali segnali interferenti intenzionali o no. Ma che relazione c'è tra un sistema autonomo e un'antenna intelligente? Le antenne intelligenti hanno bisogno di complessi algoritmi di gestione, al fine di ottenere le caratteristiche radiative desiderate. Un'alternativa a queste complesse metodologie di controllo

è l'utilizzo di algoritmi evolutivi che sono in grado di trasformare l'antenna intelligente in un sistema radiante autonomico in grado di adattare in maniera non supervisionata le proprie caratteristiche radianti alle condizioni operative dello scenario in cui deve operare.

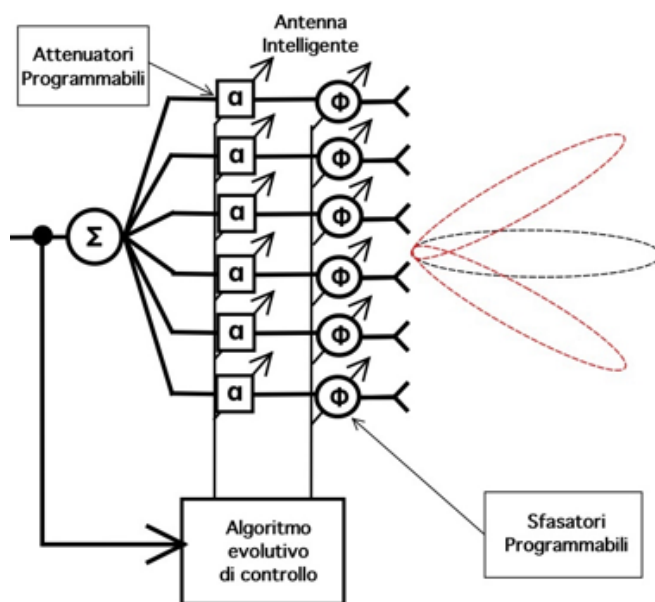


Figura 5

Schema di un'antenna a cortina intelligente, equipaggiata con un sistema di controllo autonomico basato su un algoritmo evolutivo.

L'antenna rappresentata in Figura 5 rappresenta un sistema di comunicazione autonomico, in quanto è in grado di adattarsi autonomamente alle condizioni dell'ambiente in cui andrà ad operare, in particolare tale antenna è in grado di dirigere il proprio lobo principale lungo direzioni prestabilite, eliminare segnali interferenti e incrementare il rapporto segnale rumore, mantenendo il canale wireless ad un livello ottimale. Le antenne intelligenti autonome oltre a mantenere il collegamento wireless a livelli ottimali, permettono di risparmiare energia, in quanto l'energia viene focalizzata solo nelle direzioni di interesse. I sistemi radianti riconfigurabili come quelli rappresentati in Fig. 5 sono di fondamentale importanza per i moderni sistemi autonomici, in quanto consentono di mantenere il collegamento tra le varie unità elementari in maniera ottimale.

Conclusioni

L'evoluzione dei sistemi autonomici ha portato allo sviluppo di reti fortemente pervasive e distribuite su aree di grandi dimensioni. Questa evoluzione porterà sicuramente a problematiche di comunicazione tra i nodi elementari della rete. In particolare, il sistema di comunicazione, solitamente via wireless, deve essere garantito in ogni condizione, ottimizzando le risorse energetiche degli elementi

base del sistema autonomico. In questo articolo si è proposto l'utilizzo di algoritmi ispirati al comportamento naturale per la gestione della comunicazione tra i membri elementari del sistema autonomico e l'utilizzo delle antenne intelligenti in combinazione con gli algoritmi di controllo evolutivi.

Nel lungo termine, l'applicazione delle antenne intelligenti e degli algoritmi evolutivi per la gestione della comunicazione tra gli elementi elementari dei sistemi autonomici potrebbe costituire uno dei fattori abilitanti per la realizzazione di nuovi servizi agli utenti e per l'incrementare della pervasività dei sistemi autonomici.

Riferimenti Bibliografici

- [1] IBM: *An architectural blueprint for autonomic computing*.
http://www01.ibm.com/software/tivoli/autonomic/pdfs/AC_Blueprint_White_Paper_4th.pdf (June 2006).
- [2] H.Holland, *Adaptation in Natural and Artificial Systems*, Ann Arbor, MI, University of Michigan Press, 1975. [3] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Boston, MA, Addison-Wesley, 1989.
- [4] J. Johnson and Y. Rahmat-Samii, Genetic algorithms in engineering electromagnetics, *IEEE Trans. Antennas Propag.*, vol. 39, no. 4, pp. 7- 21, 1997.
- [5] R. Haupt, "An introduction to genetic algorithms for electromagnetics," *IEEE Antennas Propag. Mag.*, vol. 37, no. 2, pp. 7-15, 1995.
- [6] P. Rocca, G. Oliveri, and A. Massa, "Differential Evolution as applied to electromagnetics," *IEEE Antennas Propag. Mag.*, vol. 53, no. 1, pp. 38- 49, Feb. 2011.
- [7] A. Qing, "A Parametric Study on Differential Evolution Based on Benchmark Electromagnetic Inverse Scattering Problem," *Proceedings 2007 IEEE Congress on Evolutionary Computation (CEC 2007)*, Tokyo, Japan, pp. 1904-1909.
- [8] Clerc, M., & Kennedy, J. (2002), The particle swarm explosion, stability, and convergence in a multidimensional complex space, *IEEE Transaction on Evolutionary Computation*, 6(1), 5873.
- [9] J. Robinson and Y. Rahmat-Samii, Particle swarm optimization in electromagnetics, *IEEE Trans. Antennas Propag.*, vol. 52, no. 2, pp. 397-407, Feb. 2004.
- [10] J. Robinson, S. Sinton, and Y. Rahmat-Samii, Particle swarm, genetic algorithm, and their hybrids: Optimization of a profiled corrugated horn antenna," in *IEEE Antennas Propagat. Soc. Int. Symp. Dig.*, vol. 1, , pp. 3143-17, 2001.
- [11] S. Caorsi, M. Donelli, A. Lommi, and A. Massa, "Location and imaging of two-dimensional scatterers by using a Particle Swarm algorithm, *Journal of Electromagnetic Waves and Applications*, vol. 18, n. 4, p. 481-494, 2004.
- [12] Donelli, M., and Massa, A., "Computational approach based on a particle swarm optimizer for microwave imaging of two-dimensional dielectric scatterers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, No. 5, 1761-1776, 2005.

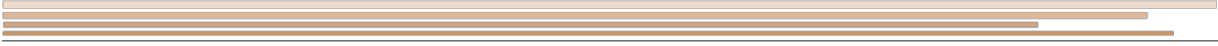
- [13] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 2, pp. 397-407, 2004.
- [14] Shyh-Jier Huang, Xian-Zong Liu, Wei-Fu Su, Ting-Chia Ou, Application of Enhanced Honey-Bee Mating Optimization Algorithm to Fault Section Estimation in Power Systems, *IEEE Transactions on Power Delivery*, vol. 28, no. 3, pp. 1944-1951, 2013.
- [15] M. Donelli, I. Craddock, D. Gibbins, and M. Sarafianou, A three-dimensional time domain microwave imaging method for breast cancer detection based on an evolutionary algorithm, *Progress in Electromagnetic Research M*, Vol. 18, pp. 179-195, 2011.
- [16] M. Dorigo, V. Maniezzo, and A. Coloni, Ant system: optimization by a colony of cooperating agents, *IEEE Trans. Syst. Man Cybern. B*, vol. 26, no. 1, pp. 294-313, 1996.
- [17] Tzuu-Hseng S. Li, Min-Chi Kao, Ping-Huan Kuo, Recognition System for Home-Service-Related Sign Language Using Entropy-Based K-Means Algorithm and ABC-Based HMM, *IEEE Transactions on Systems, Man, and Cybernetics Systems*, vol. 46, no. 1, pp. 150-162, 2015.
- [18] Nikola Todorovic, Sanja Petrovic, Bee Colony Optimization Algorithm for Nurse Rostering, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, no. 2, pp. 467-473, 2013.
- [19] Sotirios K. Goudos, Katherine Siakavara, John N. Sahalos, Novel Spiral Antenna Design Using Artificial Bee Colony Optimization for UHF RFID Applications, *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 528-531, 2014.
- [20] Lina Yang, Xu Sun, Ling Peng, Xiaojing Yao, Tianhe Chi, An Agent-Based Artificial Bee Colony (ABC) Algorithm for Hyperspectral Image Endmember Extraction in Parallel, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 10, pp. 4657-4664, 2015.
- [21] H. Banati and N. M. Bajaj, "Firefly based feature selection approach," *Int. Journal Computer Science*, Issues, 8, pp. 473-480, 2011.
- [22] M. Donelli, R. Azaro, F. De Natale, and A. Massa, "An innovative computational approach based on a particle swarm strategy for adaptive phased arrays control," *IEEE Transactions on Antennas and Propagation*, Vol. 54, no. 3, pp. 888-898, Mar. 2006.
- [23] R. Azaro, F. De Natale, E. Zeni, M. Donelli, and A. Massa, "Synthesis of a prefractal dualband monopolar antenna for GPS applications," *IEEE Antennas and Wireless Propagation Letters*, Vol. 5, no. 1, pp. 361-364, Dec. 2006.
- [24] R. Azaro, M. Donelli, M. Benedetti, P. Rocca, and A. Massa, "A GSM signals based positioning technique for mobile applications," *Microwave and Optical Technology Letters*, Vol. 50, no. 4, pp. 2128-2130, 2008. [25] S. Yang, Y. B. Gan, and A. Qing, "Sideband Suppression in Time-Modulated Linear Arrays by Differential Evolution Algorithm," *IEEE Antennas and Wireless Propagation Letters*, 1, pp. 173-175, 2002.

Biografia

Massimo Donelli, laureato in ingegneria elettronica presso l'Università degli Studi di Genova nel 1998, Ph.D. in Scienza ed Ingegneria dello Spazio presso la stessa università nel 2003. Attualmente è professore associato di campi elettromagnetici presso l'Università degli Studi di Trento, dove insegna propagazione elettromagnetica, dispositivi a microonde e tecniche di ottimizzazione. La sua attività di ricerca riguarda lo sviluppo di dispositivi e sistemi a microonde, di sistemi radianti innovativi e di applicazioni biomedicali delle microonde. E' autore di più di 200 pubblicazioni scientifiche in giornali internazionali, atti di convegni internazionali e di 7 capitoli su libro. Massimo Donelli è Senior Member dell'IEEE dal 2015 è responsabile di tutte le attività didattiche dell'European Institute of Technology presso il nodo di Trento.

Email: massimo.donelli@unitn.it





“Do you trust me?” vs “Can I trust you?” Root of Trust between Unacquainted Symbiotic Autonomous Systems

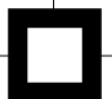
John Budenske

Abstract

The purpose of any abstract is to entice the reader to continue reading the full text of the article; there is an element of trust on the reader's part that the author(s) have properly conveyed a synopsis of the full article within the abstract. This article investigates the facets of what must exist between two symbiotic autonomous systems in order to establish a trust relationship between them (and across differing contexts). If the symbiotic autonomous systems are already acquainted or were implemented under the control of the same group of developers, then the necessary artifacts are straight forward. But, if the symbiotic autonomous systems are unacquainted (never encountered each other before), then more may be needed for trust to exist between them. If the current cyber-security state of the internet is any indication, then the two unacquainted autonomous systems will be no better off than a pair of contestants on a reality TV show. Trust is only an illusion of ambition. One cannot just assume that all encounters are with a Good Samaritan. So, if you made it this far along in this abstract, then you will enjoy the rest of this article. Trust Me!

Sommario

Lo scopo di ogni sommario è di spingere il lettore a continuare a leggere l'articolo; c'è un elemento di “fiducia” da parte del lettore che l'autore abbia correttamente condensato l'articolo nel sommario. Questo articolo esamina gli aspetti di “fiducia” che devono essere soddisfatti nella relazione tra due sistemi autonomi simbiotici per stabilire una relazione in cui ciascuno si fidi dell'altro. Se i sistemi autonomi simbiotici si conoscono o sono stati realizzati sotto il



controllo di uno stesso gruppo in genere il problema della fiducia non si pone. Tuttavia, se i sistemi autonomi simbiotici non si conoscono (non si sono mai incontrati prima) allora il problema della fiducia si pone. Se lo stato attuale della cyber-security in Internet deve essere presa ad esempio allora due sistemi autonomi che si incontrano per la prima volta sarebbero nelle condizioni di due partecipanti ad un reality show in televisione. La fiducia è solo una illusione. Uno non può attendersi che tutti gli incontri avvengano con un Buon Samaritano. L'aspettativa che può avere un sistema autonomo è analoga a quella che si può avere con un fornitore sconosciuto: questo può essere considerato affidabile nella misura in cui il contratto che si sottoscrive protegge i firmatari. Quindi, se siete arrivati a leggere fino a questo punto potete aver fiducia che il resto dell'articolo sarà di vs interesse!

Keywords: Symbiotic, Autonomous Systems, Trust, Cyber-Security, Information Assurance, Authentication, Autonomy, Intelligent Agents, Trusted-Boot.

1. Introduction

It all started when someone said to me "If something is cyber-secure, then you can trust it." And I thought "Really?" That statement is misleading from both the trust point of view as well as from the cyber-security point of view. Cyber-security and trust are not the same thing. There is a relationship between them, but it needs to be explained using terminology from something like the Risk Management Framework (RMF) [1] which has been used as a basis for cyber-security concepts and procedures. Plus, when you throw into the mix artificial intelligence, symbiotic autonomous systems, and the current state of the internet (with all its rogue actors), then you have a complicated tangle of interrelated notions across multiple domains. Thus, I started my journey down the path of thought with occasional pondering of clouds resulting in this article (which is still more editorial than artifact)

2. Symbiotic Autonomous Systems

From the point of view of artificial intelligence, autonomy, and autonomous systems, the world is quickly changing. Developers and marketers with money in their eyes are pushing new, smart and intelligent products and services into everyday life. There are almost 50 companies trying to field self-driving or autonomous automobiles [2].

We have already learned to deal with "automatic" systems around us (at least many of us have learned). Electro-mechanical systems like elevators; you push a button on the wall and it arrives; you get in and push another button and it takes you to that floor. Or an automatic cash machine; you walk up and insert a card and push buttons to authenticate your identity, and it gives you cash right out from your bank account. These are different than the autonomous systems discussed in this article in that they are performing simplified tasks in very

constrained environments. Even then, they are already intertwined into our lives; who can go a day without their smart phone?

The infusion of intelligence to create autonomous systems means that these systems will be able to behave more intelligently; perform more complicated tasks, interact with humans on a higher level of information exchange; operate in more diverse situations and environments; utilize sensors to understand what is happening around it; and can reason over what it is sensing and make acceptable if not optimal decisions on what actions to take.

Ultimately, autonomous systems will be tasked with important tasks and entrusted with important data to safeguard. They will interact with humans and become integral parts of our everyday life. Hence, they will be essentially **Symbiotic** in nature; intertwined with our physical lives, our digital lives, our actions, communications, information, and memories.

Symbiotic autonomous systems can be mobile; they can exist on a computer, within an embedded platform, or in our mobile devices. They can be referred to as intelligent systems, robots, autonomous agents, intelligent agents, etc. In this article, autonomous system will refer to all of these non-human entities that operate autonomously in our digital world, and directly or indirectly (transitively) interact with humans.

Regardless of how fast or slow these systems will infiltrate our world, it will be inevitable that they enter our lives, and there will be very, very many of them. Soon it will be impossible to go through a day without having contact with an autonomous system of some sort (autonomous taxis, autonomous calendar/schedule assistants, autonomous health providers, etc.). We already have issues with on-line automatic systems collecting and handling of personal data (e.g., phone apps that spy on us). Will we also have issues with the autonomous systems that we interact with? More importantly, the autonomy concept will place large numbers of autonomous systems together; interacting physically, electronically, virtually (e.g., envision highways and streets full of self-driving cars, or millions of personal autonomous assistants trying to buy tickets to a concert). How will autonomous systems (which humans exchange private data with) interact with other autonomous systems over that private data? Such indirect interaction could easily chain across autonomous systems, each deciding on what data to share in order to accomplish its assigned task. A single human may have direct or indirect interaction with hundreds or thousands (or even millions) of autonomous systems over the course of a day.

Internet history and the application of cyber-security as an after-thought proved to be a disaster for protecting data. My hope is that we can avoid repeating that history by getting ahead of the issues in trust and security which will arise between autonomous systems as they encounter each other in performing their assigned tasks.

3. Trust

Autonomous systems are on their way towards becoming integral in the world, and they will be symbiotic in nature and intertwined with our lives. Humans will

count on autonomous systems to perform important tasks for them, and will inevitably share private data and personal identity information (PII) with them. Symbiotic autonomous systems will need to be trustworthy to be accepted by society.

This brings us back to the statement said to me "If something is cyber-secure, then you can trust it." But, that is not completely true. It depends on the context of the trust. Just because your favorite social media website is cyber-secure (maybe its SSL1 domain certificate is authentic, or it has not been breached in recent history) does not mean you can trust it to not collect private information on your activity [3]. Being secure is in the context that the web site identity is authentic but not in the context that you can trust it to not secretly collect private data about you.

Encarta has nine definitions of trust as a noun and five as a verb. Probably the closes definition is "rely on somebody or something... to place confidence in somebody or in somebody's good qualities, especially fairness, truth, honor, or ability." [4]. Essentially, you trust the entity to do the right thing.

But that definition is a generality; it lacks context. What is "the right thing"? Do you trust "the somebody" completely and across all possible contexts, or do you constrain the trust to a single situation or context? In the case of the entity being cyber-secure, the trust could be in the context of authenticating the entity's identity, or reduction of vulnerabilities, but not necessarily in the context of trust it to do what you expect the right thing to be.

When we talk about trust in an autonomous system, there can be many contexts. You need a ride from a taxi and a self-driving automobile pulls up to give you a ride. Do we trust that the self-driving taxi will be robust? That it knows the way to your destination? That it will not crash? That it will be energy efficient? Maybe we see on the back window a little sticker that shows the self-driving taxi is a certified "Energy Star" device. Is that enough for us to decide that it can also be trusted to get us to our destination safely?

If you have taken many self-driving taxis without any trouble, you might have built up confidence and trust in taking one again. But, what if when you get in and the taxi starts up (and boots up), there is an error warning that the boot process was unable to authenticate itself such to prove that the intended correct autonomous software was properly loaded? Everything else appears OK. All the lights on the dashboard are green. Maybe a voice announces over a speaker to ignore that warning; everything is OK, and to please present your credit card before the ride instead of after. What do you trust and not trust in this scenario? Has someone switched the autonomous software? Maybe you will trust that the taxi will get you to your destination safely, but not trust that your credit card information will be protected. Or, maybe you will not trust either, and fear the taxi will dump you over a cliff and then start charging large sums of money on your credit card. Will it behave as expected or as intended? Designer, developer, implementer, owner, administrator, user; how many hands will the autonomous system pass through before I am expected to trust it? Is this also a supply chain issue?

Or, maybe the paranoia that all good cyber-security professionals harbor is escaping just a bit. My point is, when I interact with an autonomous system, I want to be able to trust it in all contexts, or at least *know* in what contexts I can trust it in. More importantly, when that autonomous system interacts with another autonomous system, I want to know the level of trust that I can expect from that interaction (and future interactions). Can I trust that my private data that I shared with the first autonomous system is not shared with the second autonomous system? And, if it must be shared to accomplish the task, will the second autonomous agent protect my private data as I wish and expect it to. Can a trust relationship be considered transitive? How do we get to the point that one autonomous system *knows* it can trust another autonomous system to do exactly what it states it is going to do (i.e., within the stated context)?

4. Cyber-Security

So, as stated above, cyber-security is not the same as trust. Cyber-security uses a risk management approach for addressing security of systems. Requirements and designs are analyzed, and vulnerabilities' risk is assessed. Risk is mitigated by adding more security controls (e.g., 2-factor authentication), etc. If we are seeking a relationship between trust and cyber-security, we could look further at risk. Cyber-security applies mechanisms and tools against vulnerabilities in order to reduce risk of system compromise. One could assume that if risk is reduced, trust will be increased.

Cyber-security provides various risk-reduction mechanisms and tools like: authentication, (passwords, certificates, etc.), integrity checking, intrusion detection/protection, malware detection, etc. Cyber-security also provides concepts like separation of access privileges, encryption to protect data, communications, and process, and attaining an *acceptable level of risk*. Risk mitigation often involves trade-offs in: processing time/memory, user constraints, development costs, customer acceptability of risk in the final implementation, etc. The risk being reduced is in "trusting something that should not have been trusted" (i.e., it's compromised).

When someone says that their system is "Cyber-secure" it means:

- it has gone through a cyber-security risk assessment and analysis, and
- the resulting level of risk (compromised/breached/etc.) is low enough to be *acceptable*
- Note that "acceptability" is very subjective! – there is no "magical state of being cyber-secure"

Cyber-security uses frameworks like NIST's Risk Management Framework (RMF) to outline areas of system design and development where vulnerabilities can exist. It provides guidance on how to assess, identify and mitigate vulnerabilities and their risk. The specific areas of vulnerability are called "security controls" and they include things like making sure that all users have low-risk user/passwords, or that the system administration makes all users change their passwords every so many days. The list of security controls becomes very specific, like a design specification or requirements list. It is a

guideline to help system developers, but if the system developer's intention is to design a smart phone app to spy on its user and collect private data, then the guideline security control of assessing whether the system properly protects sensitive data at rest using encryption is mute.

Essentially, cyber-security provides the tools that when applied correctly within a specified context, can instill trust within that same context. Authentication of an entity's identity instills trust in "the entity is who it claims to be" (in context), but that does not mean you can trust that entity with your sensitive data (outside of the intended context), unless there is some additional implicit context of trust (something like "if the entity's identity is 'John45', then the entity can be trusted with sensitive data too"). But how do we get from cyber-security tools to defining a trust relationship between autonomous systems?

Cyber-security mechanisms can be used to reduce risk and thus increase trust and once a threshold of trust has been reached within a given context of trust, then that entity can be viewed as trustworthy enough to interact with (within the given context).

5. Acquainted and Unacquainted Autonomous Systems

But, wait a minute. Society has been building networked computer systems that interact with each other for years. Financial institutions, factory automation systems, on-line purchasing all exist. Cyber-security professionals apply their knowledge to these systems, and for the most part they are cyber-secure and people trust them with their PII, and financial information, etc. Why won't that work (aside from bad guys trying to hack them)? And the answer is, yes, but remember those systems are already *acquainted*, in that they have been designed to perform only the specific actions which have been *implicitly tied* to their identity. They know which system they are to connect to and the specific protocols to follow. If the other system cannot authenticate its identity, then it is not trusted and not interacted with. But each type of connection or interaction is solely for a specific set of implicitly trusted actions; a context. If the only reason one autonomous system will connect to another autonomous systems is to perform a single type of interaction, then authenticating identity will implicitly authenticate trusting that interaction.

But this works only because the developers who implemented the computer systems and automatic machines collaborated (in some way) to apply the protocols, rules, and other cyber-security mechanisms such that the systems behave as expected. Everything does the task it is expected (agreed upon) to do. Of course, cyber criminals have proven that just authenticating an identity may not be enough to keep the transaction secure.

If some unknown entity tries to impersonate one of the two systems with stolen credentials, and they successfully scam the identity authentication step, then the impersonator could be naughty and unexpectedly steal all of the money. That is why cyber-security professionals apply multiple security controls to catch impersonators, etc. (reduce the risk, increase the trust). New tools and

approaches look at process behavior and network traffic to identify malicious actors [5].

Autonomous systems will be more "intelligent" than current automatic systems which means they will be able to do more tasks, in more complicated environments, and so the bandwidth of the interaction will be larger. When one autonomous system connects with another one that it is *unacquainted* with, it will want to know if that autonomous system is *guaranteed to be trusted to perform as expected*.

There will be many, many symbiotic autonomous systems out in the world. That means that most autonomous systems could be developed by different, distant, unfamiliar developers. We can't expect all autonomous systems to be implemented by a single development team (even with offering them time and a half overtime); and we certainly cannot expect all implementations of autonomous systems to be under the control of a single authoritarian entity that will force all autonomous systems to be implemented in the same way. Look at how horrible we have been at controlling malicious smart phone apps and phishing emails. Regardless, trusted autonomous systems will eventually need to interact with some unacquainted autonomous systems, and the need to be able to determine whether or not the unacquainted systems can be trusted will exist.

If the cyber-security state of the internet is any indication, then non-conforming, rogue, criminal, side-stepping, flim-flam autonomous systems will exist, and they will do their best to camouflage, mix-in, and hide within the conforming autonomous systems. You can't just assume that all encountered autonomous systems will be trustworthy Good Samaritans existing only to help your autonomous system. Something is needed that will allow an autonomous system to safely and trustingly interact with both acquainted and unacquainted autonomous systems.

6. Trust and Context

Earlier we stated that authentication allows trust of an entity only within the context of identity. We cannot trust that the entity will not secretly collect data on us just because we authenticate their identity unless we have also established trust within that context as well. For example, if the entity is our favorite social website "WASTEDTIME" and we established implicit trust on the company/brand WASTEDTIME that they would never pull a stunt like secretly collect data on us; then instead of authentication, the mechanism of trust is the implicit belief of loyalty for the company. The point is that "identity trust" is not the same as "private-data-protection trust"

What is of interest here is the idea of trust having a "context." The definition of trust from earlier ("to place confidence in somebody or in somebody's good qualities, especially fairness, truth, honor, or ability") is vague and abstract, and hence useless to determine whether one autonomous system can interact with another autonomous system without being duped, robbed, or flim-flammed.

The context provides an explanation or a more detailed specification of what is truly being trusted. Identity is an easy concept; you provide an entity identifying name and claim to be that entity. If a person logs onto a system as USER "Sarah" and provides the correct password to the system's credential authentication manager, then they are who they claim they are. But, that's in a perfect world where there are no bad guys trying to steal people's identity by discovering their user names and passwords. Hence, more trustworthy tools might be employed like RSA tokens and 2-Factor-Authentication (2FA).

The cyber-security profession spends a lot of resources discovering vulnerabilities in mechanisms and devising more trustworthy mechanisms and tools. For now, we leave that endeavor to them and focus on the discussion of context. More specifically, how do we recognize or define one context of trust from another?

If we go back to the dictionary, we can identify different circumstances where the word *trust* can be used. For example:

- ◆ Can I trust the **correctness** of my speedometer?
- ◆ Can I trust the **robustness** of the carnival ride (will it perform as it claims it will or crash)?
- ◆ Can I trust the **reliability** of this source of information to be truthful or is it fake-news?
- ◆ Can I trust your **integrity** with my secret password (to not share it with anyone else)?
- ◆ Can I trust your **identity** (that you are who you claim to be)?
- ◆ Can I trust your **ethics** to do the right thing?
- ◆ Can I trust the robot to **adhere** to Asimov's Three Laws of Robotics?

Is context a data-thing? a behavior-thing? a situation-thing? a law-thing? Can context be organized into a hierarchy or is it more like a semantic net? Maybe it is very distinct categories, or is it more like overlapping Venn diagrams?

I think we want to connect context to the expected result or expected behavior of an autonomous system. In systems engineering, we start with a system specification and requirements on the expected behavior of the system and its results. Maybe that is the approach we should start with? But, a system specification can be a very detailed thing, and take a long time to develop and then verify that the implemented behavior and results are in compliance with the specification.

Unfortunately, I do not have an answer for this yet. The point here is that because of differing contexts and scenarios, there needs to be a common organization, delineation, language, protocols, for recognizing and defining these differ contexts. Something like an Ontology is needed.

7. Cyber-Security and Authentication

One of the primary mechanisms used in cyber-security is authentication. It comes into play when one wants to guarantee that an entity (human, process, computer, webpage, document, etc.) is the original or actual entity that you want to interact with.

Authentication methods can utilize a number of tools like passwords, RSA tokens, digital-signatures, certificates, and other encryption-based keys. Passwords have been shown to be more riskier to count on since they can be revealed to unwanted third parties (e.g., rogue actors), but they are easier for a human to remember than an encryption key, and so cyber-security professionals devised more tools like using two-factor-authentication (2FA) where the human uses a password, and then the authenticating system sends them a token or code on their cell phone or a special device (similar to pager) to also enter. Even blockchain and distributed ledger systems might become useful authentication tools.

Some of the authentication tools utilize encryption and keys are used to verify identity, such as with digital-signatures and certificates. In order for these to work, keys must be generated and distributed such that eventually the two entities have a key that can be used to verify authentication between them. This process may include using public/private key pairs and third-party certificate servers among other tools. If authentication is to be brought into the autonomous system domain, then a similar third-party infrastructure would also be needed. Certainly, not all autonomous systems will be built by the same developers (and hence populated with each other's keys before deployment). Like web server pages, most will be developed by unacquainted developers and key distribution (similar to domain certificates) between unacquainted autonomous systems will be necessary.

For example, when you access a web site with your browser, the web site will have a certificate that has an encrypted key. Your browser will compare the certificate key with a public key (from a third-party trusted authority) to make sure the certificate is valid for the web domain you are trying to access. The certificate authenticates the identity of the web site's domain. Unfortunately, it does not do much more than that. Any bozo can claim an internet domain name and purchase/register a certificate for that domain. Does that mean you should automatically trust anything on that domain; trust anything you read on that domain; trust anything you download from that domain? Web sites are compromised all of the time and filled with malware waiting for unsuspecting downloads.

If we could define a context, like cyber-security does for identity, could we then use authentication to show that the autonomous system is trustworthy for that context. In order for this to occur, there needs to be a step where the autonomous system goes through a verification or audit process that shows it is truly trustworthy in that context, then the verification or audit process can issue the authentication certificate or affidavit.

This is already done in cyber-security. In order for a domain certificate to be issued, the owner/administrator of that domain name has to prove that they have

the required ownership during the certificate issue and registration process. For a digital-signature, the author/owner of the document will verify that it is their document before they add their digital-signature to it. Thus, part of the process for authenticating trust within a given context will be to inspect, audit and verify that the condition of the context is adhered to. The question is whether it is possible to extend authentication and its tools to other contexts.

8. Self-Authentication as a Tool of Trust

One example of a different context of trust is that of self-authentication of an embedded system. One of the vulnerabilities of embedded systems or self-booting systems (where a bad guy could gain access to the actual hardware) is to be able to determine whether or not the hardware, firmware, operating system, or other loaded software was tinkered with or compromised in some way. If an autonomous system is mobile, or it interacts with humans such that its physical system components are accessible by the public, then this could be an issue with trusting that autonomous system.

As an example, assume that you use a human nanny to watch your children. If the human nanny showed up one day while heavily intoxicated, and insisted that he was just fine to watch the kids, you wouldn't trust him to watch your children. Conversely, if you use an autonomous Artificial Intelligent nanny to watch your children, and that autonomous system has proven itself to be just as safe, robust, etc. as a human nanny, then you should be agreeable to have it watch your children when you are not home. But, what if the autonomous nanny indicated that its boot sequence detected an abnormality where it could not authenticate its loaded software to be the correct software? It may claim to be just fine since all run-time indicators are just fine. Would you trust it knowing that the system could have been tinkered with and compromised?

It is similar to the identity trust context, but it is more of a self-identity trust context. The autonomous system needs to be able to prove or certify to others (humans or autonomous systems) that it is trustworthy in the context of self-identity. Cyber-security has already started to address this kind of vulnerability. The mechanism is still authentication, but the tools are slightly different (or enhanced).

The first tool is called Trusted-Boot [6] (aka Verified-Boot or Secure-Boot). A number of chip vendors, operating system vendors, and other organizations like UEFI are supporting or providing a very similar capability which only allows signed software to run on the device [7]. It prevents the loading of BIOS, drivers, OS loaders, OS roots, OS kernels, and or application software that are not signed with an acceptable digital-signature. The specific details can vary from manufacturer to manufacturer, but it utilizes items such as EPROMS, certificates, digital-signatures, encryption, and other integrity checking tools. Often, a chain of boot steps is scripted out which involves checking the authentication of a software entity before loading it, which in turn has the keys to check the next software entity before it is loaded. If any of the authentication steps fail, the entire boot process can be stopped, and the final OS or application software will not load. The failed Trusted-Boot process can then enter an error, maintenance, and/or untrustworthy state.

The second tool is Root-of-Trust [8], [9] (aka Trust-Anchor). In order for the Trusted-Boot to work, there has to be an initial starting-point for the trust. An anchor or root is needed that upon power-up will only initiate if it has not been physically changed or tinkered. A number of approaches are being used for Root-of-Trust implementation such as EPROMS, FPGAs, and Physical Unclonable Functions (PUFs) [10]. Note that a human operator can also be the Root-of-Trust by making sure the correct initial software is booted up, but that does not help our discussion concerning autonomous systems which will be exposed to the public.

A third tool to tack on to the Trusted-Boot is to store the certificate for the identity authentication inside of the final software load that is authenticated in the Trusted-Boot chain. That way, the autonomous system cannot present its identify authentication certificate to other entities and hence it cannot be trusted by other entities unless it first Self-Authenticates itself.

Keep in mind that with this set up for self-identity context of trust, there still needs to be a third party that inspects, audits, and verifies that the Trusted-Boot chain is properly implemented along with the insertion of the identity-context certificate. Also, there needs to be a way for the third-party auditor affidavit authentication (i.e., keys) to be made available to the public. Otherwise, we are back to any bozo can claim an autonomous system identity and purchase/register a certificate.

Given these tools, one could conceptualize designing autonomous systems that can self-authenticate and then identify themselves to the other autonomous systems, hence further enhancing the identity trust that can exist between two interacting autonomous systems. And so a network of identity trusted autonomous systems could exist in which they all exchange information. But, again, that is just trust in the contexts of identity and self-authentication. How the rest of the system is implemented may be a mystery, and thus should not be trusted. If I need to trust an autonomous system with my private data or PII, then trust in those contexts will need to be defined. But, now we have a trusted state to start from. If I know how the rest of the autonomous system was implemented, etc. I can design my autonomous system to make meaningful decisions on the trustworthiness of that autonomous system once I establish identity and self-authentication trust.

9. Trust, Context, and Expected Behavior

So, from cyber-security, we have two trust contexts to think about with the authentication mechanism: identity of an entity, and self-authentication of an entity. That's a start for implementing additional trust contexts.

What I envision is that an autonomous system will be able to perform multiple tasks, and retain different types of data. All of these will result in different "behaviors" that the autonomous system will exhibit. Making an appointment at a doctor's office would be a task and how the agent executes to achieve that task is the behavior. What we are interested in as far a trusting the autonomous

system is whether the agent adheres to some properties while executing a behavior. There is an *expectation* upon that behavior

Earlier, we discussed trust having many contexts, and possibly taking a systems engineering viewpoint as context as a set of specifications and requirements upon the autonomous systems expected behavior. This would be an appealing approach since systems engineering provides very nice frameworks and tools to deal with design, implementation, verification, testing, and certifying the specifications and requirements. An alternative viewpoint would be that of Artificial Intelligence and developing Ontologies of trust, trustworthiness, contexts, expected behaviors, data protection, ethical development, etc.[11].

Since we already have a nice mechanism and tools with authentication in the contexts of identity and cyber-security is extending it to be used in the context of self-authentication, then why not keep extending it to additional contexts? All we would need to do is define another context for trust, like *Asimov's Three Laws of Robotics*[12]. If an autonomous system adheres to Asimov's Three Laws, then we can set up an authentication protocol that can be used to let other autonomous systems be able to authenticate the adherence and utilize and trust it within its interaction. Of course, the question is how does the authentication certification process know for sure that the first autonomous agent adheres to Asimov's Three Laws? It would have to be verified by an acceptable third-party; an auditor would need to inspect the autonomous system and verify that its implementation adheres.

Verification of adherence would be on the scale of performing requirements verification testing, or performing something like an International Standards Organization (ISO) inspection. Standards and specifications could be defined to allow trust contexts to be defined for a myriad of autonomous system behaviors that the public would want adherence to. Protecting and handling users on-line activity data; ability to self-drive; ability to navigate through streets, or waterways, etc.; release of PII and health data only to systems that need to know; adherence to various ethical standards like Asimov's Three Laws.

Obviously, as developers define more and more contexts of trust, then autonomous systems would most likely start to adhere to multiple contexts of trust. Possibly, the authentication of all trust contexts being adhered can be combined into a process that a single authentication tool can perform. And processes will need to be defined to handle situations like out-of-date certificates. Possibly the user is willing to take the chance and trust an autonomous system anyway and wants the option to bypass authentication (would they have a right to bypass?). The point is that a framework and infrastructure of defined contexts, auditors/inspectors, certification authorities, protocols, etc. would need to be defined from the pointed of view of certifying behaviors that the users of autonomous systems can trust (possibly include marketing and branding activities to build public trust). All of which will cost money, and where will the money come from.

Right now, governments are allowing self-driving automobiles (whose implementations are company proprietary black boxes) to practice driving on real

streets and around real people. Accidents, and unfortunately deaths, have occurred, and the question is how far does this have to escalate before a framework like authenticated trust is implemented to guarantee that symbiotic (entwined into our daily life) autonomous systems can be trusted to behave as claimed. Malicious smart phone apps have caused all sorts of mistrust with stealing logon credentials, recording private activity, and compromising PII; but it has not yet hit the threshold where people are clamoring for a similar trust context certification framework. People are happy to put up with untrustworthy apps if they are free. Will they feel the same way if trusting has a much higher risk and cost?

Conclusion

We have journeyed down the path of symbiotic autonomous systems, trust, and cyber-security. We pondered the clouds of trust having different context, and having expectations on the behaviors of autonomous systems. As we investigated the facets of what must exist between two symbiotic autonomous systems in order to establish a trust relationship between them, we only finished with questions still existing. How can we recognize and define differing contexts of trust? How can we represent trust, and trustworthiness of an entity? Is trust only an illusion? So far we have only viewed trust as a binary threshold. Should trust be represented as levels of trust that can change as an entity discovers how trustworthy the autonomous system is over time? Maybe we ARE better off not worrying about this, and follow the same paths that are being followed by smart phone apps development (would we call it "trust but verify"?). One cannot just assume that all encounters are with Good Samaritans, we must beware the unacquainted rogue autonomous systems that are out there as well. So, yes, it all started with someone saying "If something is cyber-secure, then you can trust it." And I thought "Really?"

Acknowledgements

Special thanks go to Lori Murray for her invaluable thoughts and discussions.

References

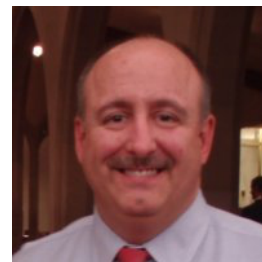
- [1] Risk Management Framework (RMF) Overview, U.S. National Institute of Standards and Technology (NIST), Computer Security Resource Center (CSRC) document, Retrieved 15 February, 2019, [[https://csrc.nist.gov/projects/risk-management/risk-management-framework-\(rmf\)-overview](https://csrc.nist.gov/projects/risk-management/risk-management-framework-(rmf)-overview)].
- [2] CBINSIGHTS, "46 Corporations Working on Autonomous Vehicles," CBINSIGHTS Research Briefs, September 4th, 2018. [<https://www.cbinsights.com/research/autonomous-driverless-vehicles-corporations-list>]
- [3] Bradbury, Danny, "Google and Facebook accused of secretly tracking users' locations," SOPHOS' Naked Security Newsletter, 25 Oct. 2018, [<https://nakedsecurity.sophos.com/2018/10/25/google-and-facebook-accused-of-secretly-tracking-users-locations>]

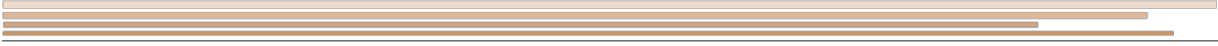
- [4] Encarta Dictionary, Encarta.com Retrieved 15 Feb., 2019, [<https://www.encarta.com>].
- [5] Murray, L., Budenske, J., Gangopadhyay, S., and Finstad R., "Cyber Resilience and Integrity Self-Awareness of Mobile Autonomous Systems," SPIE Defense+Security Symposium (SI18D), Open Architecture/Open Business Model Net-Centric Systems and Defense Transformation (Conference DS113), April 2018
- [6] Unified Extensible Firmware Interface (UEFI) Secure Boot, Wikipedia.org, Retrieved 15 February, 2019, [https://en.wikipedia.org/wiki/Unified_Extensible_Firmware_Interface#Secure_boot].
- [7] Unified Extensible Firmware Interface (UEFI) Organization, [<https://www.uefi.org>].
- [8] Zimmer, V., Krau, M., "Establishing the Root of Trust," UEFI.org document dated August 2016, Retrieved 15 February, 2019, [[https://uefi.org/sites/default/files/resources/UEFI%20RoT%20white%20paper_Final%208%208%2016%20\(003\).pdf](https://uefi.org/sites/default/files/resources/UEFI%20RoT%20white%20paper_Final%208%208%2016%20(003).pdf)]
- [9] Regenscheid, Andrew, "Roots of Trust in Mobile Devices," U.S. National Institute of Standards and Technology (NIST), Computer Security Resource Center (CSRC) document dated February 2012, Retrieved 15 February, 2019, [https://csrc.nist.gov/CSRC/media/Events/ISPAB-FEBRUARY-2012-MEETING/documents/feb1_mobility-roots-of-trust_regenscheid.pdf].
- [10] Physical Unclonable Functions, Wikipedia.org, Retrieved 15 February, 2019, [https://en.wikipedia.org/wiki/Physical_unclonable_function].
- [11] Budenske, J. R., "Expectation Verification Checking via Investigation of On-line Data Sources," Proc of the International Conference on Artificial Intelligence (IC-AI'2002), June 2002.
- [12] Asimov's Three Laws of Robotics , Wikipedia.org, Retrieved 15 February, 2019, [https://en.wikipedia.org/wiki/Three_Laws_of_Robotics].

Biography

Dr. John Budenske is a life-long renaissance technologist and entrepreneur with an interest in autonomy, cyber resiliency, robotics, and intelligent agents. He has over 30 years of R&D experience in robotics, autonomous systems, cyber security, internet of things, and human-robot interactions. His M.S. and Ph.D. were accomplished at the University of Minnesota, and he has also achieved a Certified Ethical Hacker (CEH) accreditation. He has worked for multiple large-corporation research centers and small R&D companies. He is currently employed at General Dynamics Mission Systems in Bloomington, Minnesota, USA, as a Cyber-Security and Systems Engineering Architect working on cyber-security for embedded systems.

Email: John.R.Budenske@IEEE.org





A self-flying airplane, who is ready for it?

Antonio Crespo

Abstract

The unmanned flying objects advent is a global challenging phenomenon surrounded by several uncertainties, issues and concerns related to safety, security, technology and ethics. Governments, industry and Academia are frantically dealing with various facets of the phenomenon, such as urban mobility, logistical support and integration with manned aviation. The integration of these objects into our reality involves technological and regulatory challenges, which in turn refers us to the Industry 4.0 resources and potentialities. Within such context, the present article is intended to provide a comprehensive overview of the current scenario regarding the unmanned flying objects phenomenon.

Sommario

L'arrivo di velivoli autonomi, senza pilota, porta con sé nuove sfide accompagnate da una buona dose di incertezze, sollevando aspetti legati alla sicurezza, alla tecnologia e anche all'etica. Governi, industria e accademia stanno considerando tutti questi aspetti, collocandoli nel contesto della mobilità urbana, dei supporti logistici e nella integrazione con l'aviazione classica, quella con pilota. L'integrazione di questi nuovi sistemi comporta sfide tecnologiche e nuovi inquadramenti regolamentari che a loro volta hanno diversi aspetti in comune con elementi dell'Industria 4.0 che adotta sistemi autonomi. In questo contesto l'articolo intende fornire una visione a 360 gradi dello scenario che si viene a creare con l'avvento di velivoli autonomi.

Keywords: Unmanned Flying Objects, Autonomy, Automation, RPAS, UAS, UAV, Drone, Technology, Regulation, AI, IoT, Cybersecurity

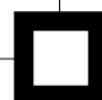
0

1

0

1

0



1. Introduction

The unmanned flying objects advent is a global challenging phenomenon. It was tabled during the World Economic Forum 2018 and motivated the realization of two international symposia within the United Nations context in that same year. Other relevant milestones include the registration by the US Federal Aviation Administration (FAA) of the millionth drone¹, the issuance by FAA of the first waiver for a fully automated beyond visual line of sight (BVLOS) operation, and the recent development of improved regulations for Unmanned Aircraft Systems (UAS), by the European Aviation Safety Agency (EASA).

Notwithstanding considered a current and hot discussion topic, it is still possible to imagine exquisite scenarios for such flying objects. You are comfortably seated in your airplane landing at San Francisco Airport via runway 28 left (28L), as the pilot announced. It is a rainy day and all you see through your window, for now, are raindrops and clouds. You keep looking and, suddenly, a big airplane pops up right beside you. Although your visibility is reduced, you cannot help but notice that the other aircraft is also landing at San Francisco, apparently headed to a parallel runway, and the two airplanes are uncomfortably close. You are finally landed and even can see the other airplane taxiing near yours. A few hours later, relaxing in your hotel room you watch in the evening news that your neighbor airplane was an all-new autonomous aircraft, moreover, the TV news made you aware that your next flight will be on that very same self-flying airplane, are you ready for it?

An apparently simple event, an airline flight, yet several uncertainties, issues and concerns come to mind: safety, security, technology, and ethics, to name a few. Situational awareness, electromagnetic environment density and security, cyber environment density and security, as well as safety and airspace capacity management are amongst the complex challenges to be addressed.

Governments, industry and Academia are frantically dealing with many facets of the unmanned flying objects phenomena, e.g. urban mobility, logistical support, integration with manned aviation, and many others, but when it comes to the regular transportation of passengers and to flying objects full autonomy, the game is undesirably restricted to few players. It is now the time to generalise this challenge, and as such, the present article is intended to provide a comprehensive overview of the current scenario regarding the unmanned flying objects phenomenon.

2. Unmanned Flying Objects, what are they?

The terminology used to define the diverse types of unmanned flying objects differs considerably and is not applied in a standardized fashion. Depending on different regions of the world, organizations, industry and criteria, such objects may be labeled as drones, unmanned aerial vehicles (UAVs), unmanned aircraft systems (UASs), remotely piloted aircraft (RPA), remotely piloted aircraft system

¹ As much as possible, the terms used throughout the article reflects the references made by the stakeholder subject to the specific presented comment or discussion; the expression "unmanned flying objects" is generally used with reference to the entire set of pilotless flying crafts regardless any technical or operational criteria.

(RPAS). However, when it comes to aircraft and airplane, the Annex 8 (Airworthiness) to the Convention on International Civil Aviation - Chicago Convention (International Civil Aviation Organization, 2018a) states:

"Aircraft. Any machine that can derive support in the atmosphere from the reactions of the air other than the reactions of the air against the earth's surface."

"Airplane. A power-driven heavier-than-air aircraft, deriving its lift in flight chiefly from aerodynamic reactions on surfaces which remain fixed under given conditions of flight."

And regarding the binomial pilot and aircraft, the Chicago Convention itself (International Civil Aviation Organization, 2006), in its Article 8 - Pilotless aircraft, brings:

"No aircraft capable of being flown without a pilot shall be flown without a pilot over the territory of a contracting State without special authorization by that State and in accordance with the terms of such authorization. Each contracting State undertakes to insure that the flight of such aircraft without a pilot in regions open to civil aircraft shall be so controlled as to obviate danger to civil aircraft."

Regardless of the various aforementioned labels used for flying objects without a pilot onboard, it is extremely important to differentiate unmanned from autonomous. When an aircraft, for instance, does not have a pilot onboard it may be labeled unmanned, and it may or may not be autonomous depending on the existence of a remote pilot.

3. Unmanned Flying Objects business: stakeholders, facts, figures and trends

Health care, social assistance, public administration, goods transportation, agriculture, construction, education, recreation, safety, security, among many others, are the sectors currently benefiting from some category of unmanned flying objects. According to the Drone Applications Report 2019 (Drone Industry Insights, 2019), these objects are being used for more than 700 different applications in several industries, whilst greatly varying in terms of size, weight, featured technology (automation-autonomy level, sensors accuracy, etc.), payload type and operational capabilities. Therefore, it is clear that unmanned flying objects business involves several stakeholders: the ordinary citizen, industry (operators, manufactures, infrastructure providers, technology-services providers), international regulatory bodies, national and local government authorities (civil aviation authority, military authority, police), Academia, etc. And the number and the inherent heterogeneity among all these stakeholders indicate how diverse the related initiatives and projects are.

In 2018 FAA registered its millionth drone (878,000 hobbyists and 122,000 commercial) and issued the second edition of its five-year roadmap addressing

the integration of UAS in the United States National Airspace System (NAS). In China, where 2.9 million UAVs were produced in 2017², there are initiatives targeting activities such as tourism, medical emergencies, urban transportation and logistics. The project EHang 184, for instance, comprises the development of an autonomous aerial vehicle aimed at medium-short distance communication and people transportation; the 200kg autonomous platform able to transport up to 100kg, for up to 25 minutes, is expected to be commercialized in 2019.

Lets take for example package delivery and air cargo: what are the most desired operation conditions by logistics stakeholders Regarding package delivery and air cargo,for these activities? Llow operating costs, larger space for goods, fewer restrictions for cargo types and the avoidance of urban traffic related delays come to mind. are amongst the most desired operation conditions by logistics stakeholdersWouldn't UAVs provide an ideal response to those needs? and it seems that UAVs might very well respond to that.

Observing a three-level (main line, a branch line and a last mile segment) intelligent logistic approach, where the final two levels are to be served by unmanned flying objects, the Chinese postal company SF Express, a 10.7 billion revenue enterprise with more than 3 billion delivered parcels in 2017, is planning on the deployment of large UAVs to cover the logistic chain between regional hubs and city hubs, and small UAVs between city hubs and business points. In 2017 the company conducted flight tests with a 1.5 ton payload UAV.

Touching urban mobility, it is well known that cities are increasingly concentrating the vast majority of the world's population, projecting a short to medium term future characterized by overly crowded streets and access roads, and inevitable counterproductive traffic bottlenecks. Due to such scenario, the giant Airbus is working on some relevant projects aimed at the inclusion of the third dimension to multimodal urban transport systems, amongst which it is worth to highlight the development of two autonomous vehicles, i.e. the single-passenger Vahana and the multi-passenger CityAirbus. Several other projects are being conducted in all continents, and they all face similar technological and regulatory challenges.

4. Technological Challenges: from automation to autonomy

The safe and secure non-segregated operation of the unmanned flying objects, considering its quantity (current and projected) and heterogeneity, requires the implementation of new functions and services, as well as the improvement and adaptation of existing ones. Registration, identification, navigation, tracking and surveillance, communication, flight planning, geo-fencing, air traffic control, air traffic flow management, air space management are amongst the functions and services requiring technology-intensive improvements. And one basic and critical premise concerning all unmanned flying object shall not be disregarded, i.e. every object is subject to contingency situations that may redound in the loss of the remote control, and as such, a certain level of autonomy, which includes

² According to Chinese industry forecasts, the UAV market (products and services) will reach more than 11 billion US dollars in 2025.

auto navigation capacity, will always be required. Within such context, the Industry 4.0 reality decisively contributes to shaping the way manned, unmanned and autonomous aviation is being conceived and integrated.

Effective and reliable detection and avoidance (DAA) capabilities are essential for non-segregated operations. Specifically regarding RPAS, the required communication performance is of particular importance, where signal latency between the remote pilot and the air traffic services units may jeopardize airspace safety. The command and control (C2) link robustness, availability, reliability shall ensure safe and secure operations by avoiding interference (anti-jamming), hijacking (anti-intrusion). In summary, the unmanned flying objects embarked technology shall ensure they interact with the ATM ecosystem just like the manned aircraft, and therefore fulfill proper communication, navigation and surveillance requirements.

Moreover, nowadays a relevant portion of all unmanned flying objects features a high level of automation, and so is the case for the state of the art manned aircraft, where current capabilities include auto flight and navigation, auto land, thrust management, systems management, health monitoring and reporting. Such increasing automation levels might for sure be seen as steps towards autonomy.

The diverse levels of automation featured in such objects versus the ideal automation condition required for a specific type of operation or for an airspace class³ also represents an important challenge, and its complexity exponentially increases as the automation levels converge to full autonomy. According to the US Department of Defense (DoD), autonomy is defined as (Department of Defense, 2018):

"...the computational capability for intelligent behavior that can perform complex missions in challenging environments with greatly reduced need for human intervention, while promoting effective man-machine interaction."

and its technological implementation challenges are grouped in four key areas, i.e.: a) machine perception, reasoning and intelligence; b) human/autonomous system interaction and collaboration; c) scalable teaming of autonomous systems; and d) test, evaluation, validation and verification. The technological evolution within such key areas, as well as its effective availability and implementation drive the autonomy evolutionary scenario, as presented in table 1.

³ As prescribed by the Annex 11 (Air Traffic Services) to the Chicago Convention (International Civil Aviation Organization, 2018b) In the ATM ecosystem, and for Air Traffic Services provision purposes, the airspace is divided in seven classes (A, B, C, D, E, F, G); the classification is determined according to the permitted type of operation (IFR, VFR), the provided services and the way flights separations are accomplished.

Autonomy Level	Human-System Interaction	Binomial Aircraft-Crew Scenario	Scenario Description
1	Human Operated	Assisted Crew (low level)	A human operator makes the decisions. The system has no autonomous control of its environment although it may have information-only responses to sensed data.
2	Human Delegated	Assisted Crew (high level)	The system can perform many functions independently of human control when delegated to do so. This level embraces some automatic controls (e.g. engine control) and other low level automation that must be activated / deactivated by human input and must act in mutual exclusion of human operation.
3	Human Supervised	Reduced Crew	The system can perform a wide variety of activities when given top-level permissions or direction by a human. Both the human and the system can initiate behaviors based on sensed data, but the system can do so only if within the scope of its currently directed tasks.
4	Full Autonomy	Crewless	The system receives goals from humans and translates them into tasks to be performed without human interactions. A human could still be in the loop for emergency situations or goals changing, although in practice there may be significant time delays before any human intervention occurs.

Table 1

Autonomy evolutionary scenario and the binomial aircraft-crew (based on the U.S. DoD proposed framework)

Within the aviation domain, autonomous systems shall be featured with technology and resources enabling them to independently (from human interference) accomplish complex missions in complex environments, keeping safety and security as paramount operation conditions. As such, the autonomous conduction of complex flights requires connectivity, accurate sensing, real-time communication, real-time big data processing and real-time decision-making (perceives, learns, abstract, reason, plan and decide).

Although it is clear that the (deciding) intelligent autonomous agent piloting the flying object is supposed to manage all the interactions inherent to the ATM ecosystem, it is unclear whether AI systems will happen to be certifiable according to manned aviation airworthiness standards, considering that the acceptable probability of software / hardware catastrophic failure (fatal injury or incapacitation) for commercial passengers air transportation is of 1 accident per 10^9 flight-hours. Therefore, on top of the technological challenges, some regulatory challenges shall also have to be considered.

5. Regulation Challenges and Efforts

According to IATA (International Air Transport Association, 2016), between January 2013 and August 2015, 856 safety related documents reported the existence of drones flying dangerously close to aircraft and airports worldwide. Transport Canada registered 148 drone incidents in 2016, against 41 in 2014. Between 19 and 21 December 2018, Gatwick airport faced a huge operations disruption due to drones' interference, which affected more than one thousand flights and 140,000 passengers. In January 8th, 2019, Heathrow airport had its operations suspended for more than one hour for the same reason.

Although the interests of all stakeholders are not expected to be fully aligned or convergent, harmonization and interoperability are key concepts underlying the safe and secure utilization of such promising flying tools. Moreover, the integration of such objects with existing manned aviation ecosystem is of paramount importance, and the implementation of a proper regulatory framework is a key step towards overcoming such challenge.

The harmonized and integrated operations of all unmanned flying objects require regulations framing its design and manufacture, the airspace management and access, and for sure certainly the interaction with manned aircraft. Therefore, some extremely important existing aviation regulation principles shall still apply, i.e.: a) the development and application of airworthiness standards; b) the definition and enforcement of operational rules; c) the requirement for educated operators and remote pilots; and d) the establishment of safety and security management frameworks.

As mentioned previously, such objects are of several types, with varying sizes and weights, and are being used differently among several industry sectors, and on top of that they are expected to transport human beings anytime soon in the near future. Such outlook poses huge pressure on many civil aviation authorities and other regulatory bodies worldwide. The regulating initiatives currently in progress are facing complex challenges. i.e.:

- proper and timely object / operator identification;
- surveillance and real-time communication requirements;
- airworthiness:
 - manned aviation airworthiness standards application;
 - aviation grade certification of originally 'non-aviation-related technologies', e.g. AI, IoT, 5G, cloud platforms, etc;
- operations standardization;
- states heterogeneity in terms of regulatory and technological scenarios;
- the requirement for industry to share with regulators accurate and up-to-date information on products, embarked technologies and services.

In order to address some of these challenges, the International Civil Aviation Organization (ICAO) is currently reviewing some very important SARPs references, such as the Annexes 1 (Personnel Licensing), 2 (Rules of the Air), 6 (Operation of Aircraft), 8 (Airworthiness), 10 (Aeronautical Telecommunications),

13 (Aircraft Accident and Incident Investigation) and 18 (Safe Transport of Dangerous Goods) to the Chicago Convention. The reviewing process of Annex 6 seeks to add a document part specifically addressing RPAS operations, which is expected to be effective by 2020. The document prescribes the implementation of a risk-based three-category (A, B, C) operations taxonomy.

Category A (low risk) operations are characterized by visual-line-of-sight (VLOS) with the flying object, which shall maintain specified distances from aerodromes and people and observe a maximum flying height above ground level. These operations could be exempted from any specific authorization and the flying objects itself not subject to ICAO standards and recommended practices (SARPs).

Category B would embrace medium risk operations subject to regulations aimed at reducing its risk to a low level. The category would include operations with potential risks to persons, and with the possibility of sharing the airspace with other aircraft, at risk levels lower than those requiring traditional aviation certifications. Category B operations would probably be restricted according to the object type and operating personnel, as well as require approvals based on risk assessment and mitigations; an authorization process under government oversight is expected to apply.

High risk RPAS operations in the category C are recognized as similar to manned aircraft operations and are expected to abide by traditional regulating methods, inby which certifications would be required for the operator (type of operation), the flight crew (license) and the flying object itself (airworthiness), in accordance with ICAO SARPs.

ICAO envisages a full RPAS integration into the ATM framework, where they would be regularly issued the applicable certifications. Particularly, remote pilots would be granted licenses issued based on a competency framework⁴ very similar to the one currently applied to traditional pilots. It is worth to highlighting though that the ICAO current regulatory efforts scope excludes both the transportation of persons and fully autonomous aircraft.

The European Aviation Safety Agency (EASA) recently implemented a performance/risk-based regulation aiming at framing UAS operations. According to such regulatory framework an UAS shall be categorized as either open, specific or certified. The 'open' category includes objects (not heavier than 25kg) bound to perform low risk VLOS operations limited to heights not exceeding 120 meters above the surface. Some UAS basic design requirements might be enforced by means of the regular European CE-marking certification process.

The objects in category 'specific' are expected to be involved in operation conditions deemed as of increased risk, what which would require proper authorization based on specific operation risk assessment (SORA). User of UAS under this category may optionally be granted an operator approval subject to the definition of standard operational scenarios. The 'certified' category frames

⁴ The competency framework for remote pilots is suggested to include competencies such as problem solving and decision-making, management of non-normal situations, leadership, teamwork, workload management, automation management, situation awareness, among others.

operations falling under any of the following conditions: a) usage of an object with dimensions above 3 meters or kinetic energy above 34 kJ, and intended to fly over open assemblies of people; b) people transportation; c) dangerous goods carriage potentially posing high risk for third parties in case of accident. Similarly to manned aviation, operations of UAS included in such category require certifications for both the object and the operator, as well as pilot licensing for non-autonomous flights.

The FAA joint technologically driven and regulatory efforts towards UAS full integration rely on a scalable matrix encompassing the requirements for seven complexity-increasing operational capabilities, i.e. (Federal Aviation Administration, 2018):

- A) operations over people - small UAS operating over people not directly involved with the operation;
- B) expanded operations - small UAS beyond visual line of sight operations, swarms⁵ and on-airport operations;
- C) small UAS package delivery operations - small UAS fleet accomplishing external load and agricultural operations involving multiple launches and landings for materials transportation;
- D) non-segregated operations - restricted UAS operations in controlled airspace coexisting with manned aviation, including properly equipped large UAS flying under IFR and according to submitted flight plans; e.g. interstate delivery and small cargo operations;
- E) scheduled operations - regular scheduled UAS arrivals/departures at airports class B, C and D and large cargo operations with RPAs, including IFR operations and subject to flying object capabilities and ATC infrastructure;
- F) large carriers logistics operations - cargo transportation with RPAs in domestic airspace; and
- G) passenger transportation operations - air taxi services with remote pilots conditional to vehicle performance requirements, certification type, featured equipment and automation level.

Notwithstanding the aforementioned relevant regulating initiatives, the current global situation reflects a scenario in which the majority of the States are critically ineffective regarding unmanned flying objects registration, which implies unreachable operators. The lack of awareness on the objects population (number, types, operations, etc.) redounds in difficulties to dimension, plan and implement the required supporting infrastructure and regulatory framework.

⁵ *Swarming drones form a set of flying objects able to communicate with each other and to operate in a coordinated way.*

Consequentially effective UAS international IFR⁶ operations regulation is not expected to be in force before 2024.

6. Outlook

In a fully harmonized and integrated scenario, remotely piloted and autonomous flying objects shall be able to effectively interact (sense, perceive, analyze, communicate, plan, decide and act) with all systems inherent to the ATM ecosystem. Hence, international organizations, Academia and industry are working together in the development and implementation of new approaches and technologies aimed at coping with the rapidly growing number, types and applications of unmanned flying objects. And when also considering the current growth rates of manned aviation, the safe, secure and efficient non-segregated utilization of airspace is expected to become a challenge with ever increasing complexity.

NASA is currently one of the most active stakeholders working towards the UAS integration in non-segregated airspace and considering autonomous systems for cargo transportation, covering technical challenges related to command and control, detection and avoidance, and systems integration and operationalization. NASA's vision foresees manned and unmanned aircraft simultaneously and routinely operating through all flight phases in the U.S. national airspace, considered specific air traffic management requirements and system performance capabilities. And such This vision is supported by the four pillars and related enablers presented in table 2. (National Aeronautics and Space Administration, 2018)

NASA's Vision Pillars	UAS Technologies	Air Traffic Management (ATM) Services and Infrastructure	Operational Regulations, Policies and Guidelines	Public Acceptance and Trust
Enablers	Airport Operations Technology	Airport Infrastructure	ATM Regulations, Policies and Guidelines	Cyber security Criteria and Methods of Compliance
	Airworthiness Standards	ATM Infrastructure	Airworthiness Regulations, Policies and Guidelines	Legal and Privacy Rules
	Command, Control and Communication Technologies	Non-FAA Managed Airspace Infrastructure	Operating Rules, Regulations and Procedures	Noise Reductions
	Detect and Avoid (DAA)	Radio Frequency (RF) Spectrum Management	Safety Risk Management and Methods of Compliance	Physical Security Criteria and Methods of Compliance

⁶ IFR is the acronym used to designate instrument flight rules, and IFR operations are those conducted in accordance with the instrument flight rules, as prescribed by the Annex 2 (Rules of the Air) to the Chicago Convention (International Civil Aviation Organization, 2005). IFR is defined in opposition VFR (Visual Flight Rules), and IFR operations related certifications are more stringent than the VFR ones.

	Flight and Health Management Systems	Test Ranges and Simulation Facilities		Public Safety Confidence
	Ground Control Station (GCS) Technologies			
	Hazard Avoidance			
	Highly Automated Architectures			
	Navigation			
	Power and Propulsion			
	Weather Awareness and Avoidance			

Table 2
NASA vision on UAS integration: pillars and enablers

Additionally, the so called UAS Traffic Management (UTM) initiative seeks a globally harmonized framework able to cope with the challenge of the safe and secure integration of these flying objects into domestic airspace. The UTM refers to a set of services and an all-encompassing framework (specifically tailored airspace planning, adapted communication-navigation-surveillance infrastructure and proper regulation) designed to manage multiple UAS operations, as well as to ensure interoperability with the existing ATM ecosystem. The UTM concept of operations proposes the provision of cooperative traffic management for all objects flying within a designated operational volume (conditional to technical and operational requirements, and airspace class), and describes its architecture, actors, roles and responsibilities. Important aviation stakeholders such as the European Organization for the Safety of Air Navigation (EUROCONTROL) envisage the UTM as part of the ATM, where the service providers would have some delegated responsibilities.

Industry 4.0 technologies are deemed as key UTM enablers. Some relevant initiatives are focusing on the development of cloud platforms aimed at supporting UTM implementation. The goal is to provide service oriented UTM related data exchange (identification, flight information, weather information, communication, surveillance and tracking, air traffic services information and data) accessible by all interested stakeholders. Other enterprises in China are working on the implementation of a 'space-based internet' concept expected to overcome the coverage limitations of the regular data transmission means established with UAVs (mobile communication or line of sight microwave) in a UTM environment. Flight risk assessment relying on big data (flight plan load, aircraft type, flying object type, mission type, location, etc.) multi-intelligence fusion is also considered for the improvement of the integrated manned-unmanned operations.

5G technology is expected to revolutionize the connectivity worldwide, by enabling ultra-reliable low latency communications ($\pm 0.5\text{ms}$ for objects moving

at up to 500 km/h) through enhanced mobile broadband (greater data-bandwidth) for any massive machine type communications scenario ($\pm 10^6$ connections/km²) with AI support for automatic elimination of interference; IoT boosted by 5G technology is foreseen as one of the key UTM enablers.

Likewise, AI is one of the key enablers of autonomy in the aviation domain. AI algorithms designed to support real-time decision-making and sensing processes combined with Robotics might conform the backbone of the features required to manage the complex interactions to be in place with amongst all aviation related systems, services and agents, e.g. other flying objects, airports, Air Traffic Control system, Aeronautical Information system, Aeronautical Meteorology system, etc. Machine learning and deep learning capability to deal with image processing may improve the efficacy of DAA systems. However, the certification of AI systems is one of the most challenging aspects towards the full integration of autonomous flying objects into the aviation regular ecosystem.

While UTM effective implementation, AI systems proper certification and full integration with ATM are not achieved, the design and application of anti-UAS measures are of extreme importance. In contingency situations, continuous monitoring of UAS existence and activities shall be combined with effective countermeasures in order to prevent undesirable interference on manned aviation operations. Anti-UAS measures may include radio-frequency (RF) signal analyzers, RF direction finders, jammers and blockers, anti-UAS RADAR based systems, optical surveillance devices, and means of interception.

Finally, in addition to the uncertainties related to UTM implementation and AI certification, the aviation community perceives some relevant approach gaps regarding the unmanned flying objects phenomenon in general; for instance, what would be (is) the impact of one million flying cameras on over people's privacy? And how comfortable would you be knowing that autonomous air taxis (carrying who?) are flying over your house. Or even, is the triad autonomy-electromagnetic environment security-cyber security being properly articulated? These are some gaps that need to be covered globally, but before answering these questions, it would be extremely important for all 'phenomenon' stakeholders to agree on common visions, languages and references about the future of such objects. Then, there would be only one question left: how would you feel about being informed by the news on your TV that your next flight is pilotless? Buckle up, it is time to takeoff into the future!

References

Department of Defense. 2018. DoD Autonomy Roadmap. Retrieved Feb 4, 2019, from <https://ndiastorage.blob.core.usgovcloudapi.net/ndia/2018/science/Kearns.pdf>.

Drone Industry Insights. 2019. The Drone Applications Report 2019. Retrieved Feb 4, 2019, from <https://www.droneii.com/drone-applications-2019#1525106654181-a2b63cd6-e0c3>.

Federal Aviation Administration. 2018. Unmanned Aircraft System (UAS) Traffic Management (UTM). Retrieved Feb 4, 2019, from <https://utm.arc.nasa.gov/docs/2018-UTM-ConOps-v1.0.pdf>.

International Air Transport Association. 2016. Safety Awareness for Users of Remotely Piloted Aircraft (RPA) in Close Vicinity of Airports. Retrieved Jan 31, 2019, from <https://www.iata.org/whatwedo/safety/Documents/Safety-Awareness-for-RPA-users-joint-statement>.

International Civil Aviation Organization. 2005. Annex 2 to the Convention on International Civil Aviation: Rules of the Air. Montreal, 10ed, July 2006.

International Civil Aviation Organization. 2006. Convention on International Civil Aviation. Montreal, 9ed, July 2006.

International Civil Aviation Organization. 2018a. Annex 8 to the Convention on International Civil Aviation: Airworthiness of Aircraft. Montreal, 12ed, July 2018.

International Civil Aviation Organization. 2018b. Annex 11 to the Convention on International Civil Aviation: Air Traffic Services. Montreal, 15ed, July 2018.

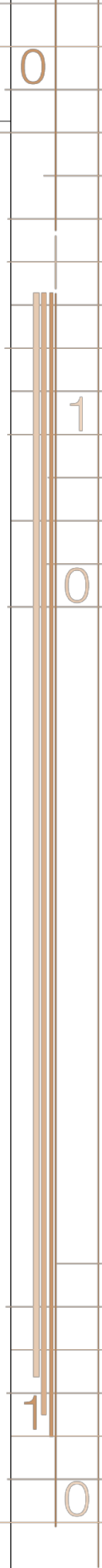
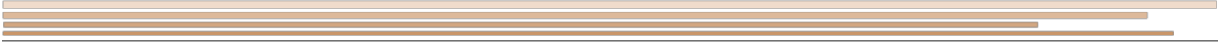
National Aeronautics and Space Administration. 2018. Unmanned Aircraft Systems (UAS) Integration in the National Airspace System (NAS) Project. Retrieved Jan 29th, 2019, from <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20170006091.pdf>.

Biografia

Antonio Crespo, Bachelor of Aeronautical Sciences by the Air Force Academy and Bachelor of Social Sciences (Sociology and Political Sciences) by the Federal University of Santa Catarina. Postgraduate Diploma in Electromagnetic Environment Analysis by the Electronic Engineering Division of the Aeronautics Institute of Technology. Master (MSc) in Computer Science by the University of Brasilia. He served as Executive Secretary of the Commission for the Brazilian Airspace Control System Preparation for the FIFA 2014 World Cup, as a member of the Special Operations Committee of the National Commission of Airports Authorities, and as Deputy-coordinator of the the Olympics 2016 Air Transportation Operations Command and Control Center. Former Air Navigation Commissioner and Alternate Representative of Brazil to the International Civil Aviation Organization (ICAO) Council, Head of the Operations Planning Division and Executive Secretary of the Project for Optimization, Restructuring and Operational Innovation Committee in the Brazilian Department of Airspace Control. Currently, CEO of the DreamerGate Inc. and researcher in the Information and Systems Engineering PhD program at Concordia University.

Email: antonio.crespo@mail.concordia.ca





Defense Machines: Digitization, Cybernetics and Autonomous Network Security

Aman Singh

Abstract

The information age has transformed our society into a global self-organizing network of autonomous digital systems. This global network is continuously growing with millions of new heterogeneous devices joining cyberspace. The growth is fueled by advancements in both hardware and software technologies such as silicon photonics, cloud computing, internet of things and artificial intelligence. As more devices enter cyberspace, they also increase the attack surface and ultimately rendering the entire network vulnerable. Management of these systems by humans will not be feasible due to massive scale, heterogeneity, and software complexity. For a healthy cyberspace, we also need autonomous defense systems that can learn, adapt and defend against emerging threats. Using the principles of cybernetics, we present an autonomous network security operations model. We apply this model to three different kinds of environment. First, the model is applied to secure network operations of a smart home environment. Second, it is applied to defend against disinformation spread in online social network services, and finally, it is applied to model software code itself to create self-aware systems.

Sommario

L'età dell'Informazione ha trasformato la nostra società in una rete globale auto-configurante composta da sistemi digitali autonomi. Questa rete globale continua a crescere con milioni di nuove entità eterogenee che nascono nel cyberspace. Questa crescita è alimentata dai progressi tecnologici, sia nell'hardware sia nel software, quali quelli nella fotonica su silicio, cloud computing, IoT e intelligenza artificiale. La crescente presenza di nuove entità nel cyberspace aumenta anche l'area suscettibile ad attacchi rendendo l'intera rete vulnerabile. La gestione di questi sistemi da parte di persone sarà sempre meno fattibile visti i volumi crescenti,



l'eterogeneità e la complessità che ne deriva. L'obiettivo di avere un cyberspace sicuro potrà essere perseguito solo avvalendosi di sistemi di difesa autonomi in grado di rilevare e contrastare ogni minaccia. Utilizzando i principi della cibernetica, presento in questo articolo un modello eseguibile di un sistema di rete sicura autonomo. Come esempio applicativo questo modello viene applicato nel contesto di una rete sicura per un ambiente domestico smart. Inoltre, evidenzio come lo stesso modello possa essere applicato per contrastare tentativi di disinformazione e per proteggere il codice software.

Keywords: Cybersecurity, Autonomous Defense, Cybernetics, Information Spaces, Network Modeling, Monitoring and Management, Disinformation Diffusion and Defense, Self-aware Software

1. Introduction

The introduction of the general-purpose Electronic Numerical Integrator and Computer (ENIAC) [1] in 1946 and the mathematical foundations of communication in 1948 [2] marked the arrival of the digital information age. It has put us on an innovation path never seen before in human history. With electrons performing computation, photons in-charge of transport and ions taking care of storage needs, the global digital information capacity will continue to increase with improved materials and computing principles. In the last decade, cloud computing, Internet of Things (IoT) and advances in artificial intelligence methods have accelerated digitization of every aspect of our society. Furthermore, the digital transformation initiatives such as the fourth industrial revolution [3] and Society 5.0 [4] in Japan provide a structured growth path. The major factor of transformation has been the Internet which has become the nervous system of our society. It has created a single unified digital communication fabric where anything can be controlled with a single button from anywhere on the globe. Without the Internet, our modern society would come to a halt. In order to adequately secure this complex network against emerging threats, better system modeling and design principles are essential.

The 1940s also witnessed Macy conferences in New York [5] that brought together multi-disciplinary scientists to understand common foundational principles of information processing in different fields of science. The common principles resulted in Cybernetics meta-discipline that describes control and communication in any system capable of receiving, storing and processing information [6]. With systems and control theories at its core, cybernetics helps us understand and define the functions and processes of a system that have goals and participate in action-sense-evaluate cycles. Any closed network system can be modeled using cybernetics principles. In particular, second-order cybernetics [7] is useful in modeling an information system by including an external observer that can monitor & control system environment. A useful

property of cybernetics is the recursive reduction of systems, for example, the Internet is a group of heterogeneous networks, networks provide services, services run on machines, machines run programs, programs run functions and so on; the same abstract structural control principles can be applied to every subsystem in this information flow chain.

In the rest of the paper, Section 2 describes the information abstract spaces around us, and how the cyberspace growth has been accelerating due to recent advancements in system virtualization, IoT, and AI technologies. Section 3 describes the dual nature of technology and the need for autonomous network defense. Section 4 describes the security threat actors and attack life-cycle. In Section 5, using cybernetics principles, we describe a network operations model with autonomous defense agents. We apply the same model for securing online social networks communications against disinformation. Furthermore, the model is applied to enhance software code security resulting in self-aware systems. Finally, we conclude with the need for a closer collaboration between academia, industry, and government bodies in order to develop better technologies to improve the safety and security of cyberspace.

2. Information Spaces and Digitization

For better understanding the digitization of our society, we can abstract all information into three spaces around us – physical, societal, and cyber. The physical-space is the foundation that constitutes all living and non-living things around us. The societal-space consists of dynamic networks that operate our cities, nations and ultimately our civilization, for example, healthcare, education, agriculture, transportation, commerce, finance, government, law, and politics. The cyberspace consists of digital representations of objects from all information spaces including itself, and facilitates interconnection, communication, and control. The representation of an object in cyberspace is called a digital twin that has an identity along with the object's behavioral patterns, state, and control functions [9]. For example, digitizing a business will involve modeling business users, entities and processes into digital twin objects with associated properties and behavior. The business operations are executed with digital twin interactions that perform everyday tasks. Figure 1 illustrates the three abstract information spaces and the corresponding interactions, we can also observe with growing digitization, the cyberspace is becoming the default bidirectional proxy for interaction between the societal and physical spaces.

Software is eating the world [40], and the recent rush in the appetite can be traced back to the trinity of - Virtualization, IoT and AI technologies. Virtualization has enabled cheap, secure and location-agnostic computing in cloud, edge, and even mobile phones. Software-defined-Infrastructure (SDI) [8] automates the provisioning of compute, network and storage resources based upon service requirements across location boundaries. Open source software has enabled rapid innovation in application development.

A network system can be abstracted into a set of service processes that work together to accomplish well-defined tasks. For example, a modern city can be modeled as a network computing system providing diverse services such as

water, waste, electricity, gas, heating, public transport, and traffic. For control and optimization of these services, they are digitized using sensors and actuators, Things, that enable integration of the physical-space with the cyberspace. In general, these IoT systems contain three different phases - collection, transmission, and processing. The collection phase captures system events, the transmission phase communicates the events to a system controller, and the processing phase generates system output actions based upon input events. This real-time feedback provides control and automation in the networks.

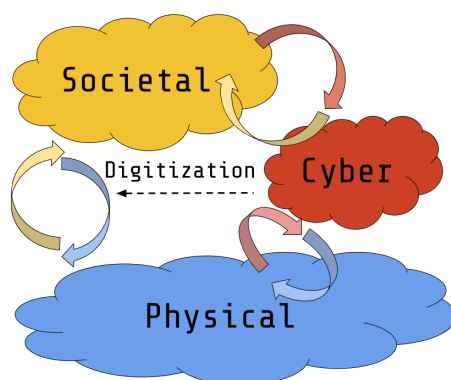


Figure 1
Information spaces and Digitization

Since the inception of computing machines, humans have been trying to replicate intelligence [10]. The initial decades saw slow growth, but in recent years there have been significant new and improved machine learning methods proposed. With the reduced cost and increased capacity in compute, network and storage technologies, we have started to capture raw system event data. The massive raw data combined with improvements in machine learning methods have shown breakthroughs in natural language processing, computer vision, speech recognition, and pattern recognition tasks. The machines are gradually gaining all five senses - vision, hearing, touch, smell, and taste. With enough available contextual data, the machines can now learn the inherent structure of process operations and automate workflows, perform pattern recognition, detect objects, translate text, write songs, draw paintings, drive cars and have human-like conversations.

Figure 2 illustrates the data, control and semantic planes in a system. The control plane controls “how” the data flows in the system. The addition of an AI layer to an existing system provides semantic context information that improves the control by learning “what” data is flowing, and “why” the system behaves as observed. In other words, it improves upon the baseline control assumptions of the system by learning from process observations. This helps in optimization, anomaly detections, and state predictions. These context-aware systems create a symbiotic relationship between humans and machines. Currently, we are observing the first wave of AI enabled systems that automate and optimize business processes and some human body parts. The next wave will enable

symbiosis in societal networks, and finally, when an AI system starts to learn from itself, the singularity will be achieved [11]. However, we are still a few decades away from these goals, as the current AI systems work like black magic boxes. For establishing confidence and trust in AI systems, they have to be accountable in explaining why a decision was made, and transparent in describing and reproducing decision process. The Explainable AI (XAI) [12] DARPA program is looking to address these areas by producing learning models that human users can understand, trust and effectively manage.

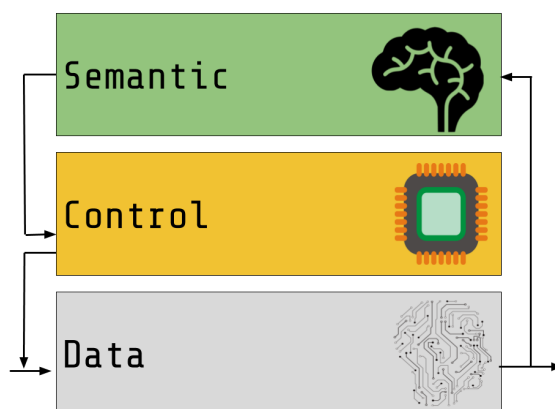


Figure 2
System Information Planes

3. Technology Duality

The technological progress comes at a price. It can be used for both good and bad purposes. While making the world more connected, the cyberspace has also enabled billions of robocalls, email spams, mass surveillance using Closed-Circuit Television (CCTV), accelerated diffusion of fake news through social media, malware, and advance military drones. There is an inherent duality in technology having both offensive and defensive behaviors that are in a constant battle of system control. This behavior equilibrium provides the necessary impetus for system growth. The cyberspace equilibrium is being disrupted with the accelerated introduction of millions of new vulnerable devices every day and offensive AI methods used to launch more sophisticated attacks.

The software running our societal networks is ultimately developed by humans, coders, that often introduces programming code errors resulting in vulnerable systems. On any regular day, Internet devices, hosts, are bombarded with attacks, most of them from automated software, bots, scanning for vulnerabilities. Furthermore, human users of these systems themselves are threats if they are not trained properly, for example, phishing attacks. Everyday news about security data breaches have become common because of these inherent software and human vulnerabilities. If we can take a landscape picture of Internet cyberspace, it will be a toxic wasteland of bits with occasional heavily fortified citadels.

Most consumer IoT devices that have limited functionality contain vulnerable software, insecure default configuration, and no security management functions. They increase the attack surface, rendering the whole system vulnerable. A compromised IoT system can be used for Denial-of-Service (DoS) attacks, digital crime proxies, ransomware, data theft, and more recently, cryptocurrency mining cartels [13]. For example, the recent Mirai worm targeted vulnerable connected consumer IoT devices such as camera, baby monitors and switches, these devices became part of massive DoS attack campaigns [14]. With AI, the complexity of security attacks has increased enormously with more attacks being autonomous [15]. Once an attacker assigns an objective, the autonomous software bots continuously scan for vulnerable machines and humans on online sites. When a vulnerable target is identified, automated attack plans are dynamically realized without any human interventions. The cyberspace has now become the fifth domain of warfare after land, sea, air, and space.

4. Cyberspace Threat Actors & Attack Life-Cycle

The cyberspace threat actors can be broadly classified into four major categories – criminals, hacktivists, nation states, and insiders. The first three actors reside outside a network system operational perimeter. The cyber criminals are mostly interested in monetary incentives. The cyber hacktivists target specific entities with a goal of data theft or simply vandalism to tarnish an organization's reputation. The nation-state actors also target specific entities especially foreign governments and corporations with a goal of espionage, intellectual property theft, information manipulation, and destruction. The insider actors can cause the most damage to a system. The major risk is due to access policies, as these actors can turn rogue anytime. They are mostly employees or third-party contractors that are looking for revenge, profit gains or are under external pressure. For targeted entities, nation-states can also introduce persistent hardware or software implants via supply chain vulnerabilities [16] that allow zero-day access.

For an outsider, an attack life-cycle follows a standard set of steps, defined as intrusion-kill-chain, for gaining full access to remote systems [17]. The life-cycle steps are - reconnaissance, weaponization, delivery, exploitation, installation, command & control, and objective actions. The reconnaissance step involves research and selection of targets using digital footprints such as emails, social profiles, mailing lists, and conference proceedings. Weaponization involves remote access malware packaged with vulnerability exploits as payload files such as Flash video, PDF document, Office documents, audio or image files. The delivery step involves the transmission of the weaponized payload via different channels such as email attachments, physical USB drives, website URL access or open wireless networks. The exploitation step involves triggering the application code vulnerability for successful exploits. This step infects most systems that are not patched or the exploit is a zero-day threat. The installation step hides the backdoor in the system and makes it persistent. The command & control step involves communication with external malicious servers that provides further mission actions to compromised network hosts. Finally, the

objective actions are achieved that can include data theft, destruction, extortion, or intrusion of another network.

We have seen different types of threat actors; their motivation and how cyber-attacks are performed. As discussed previously, with the addition of AI, the cyber attacks are becoming more sophisticated. The malicious use of AI report [15] highlights these attacks on all three information spaces – cyber, physical, and societal. In the cyberspace, systems will experience an increase in social engineering attacks with human-level interactions and phishing attack campaigns; automated code vulnerability discovery and exploit generation; human-like service access for DoS; and finally, bot vs. bot warfare will be seen. In the physical-space, commercial high-skill weapons with low skill humans; fully autonomous weapons; and finally, robot vs. robot warfare will be seen. In the societal-space, as the motive is to disrupt state equilibrium for system control, we can anticipate mass surveillance of citizens; massive fake news generation and diffusion channels, for example, deepfakes [18]; personalized disinformation campaigns; denial-of-information attacks; and finally, political unrests with human vs. human conflicts.

5. Autonomous Network Defense

How can we defend our systems? As all the threats originate from programming code, establishing trust in our computing infrastructure, learning from operations, designing a layered defense architecture, taking proactive measures and adapting to emerging threats should provide sufficient coverage. But, most of these design and operations best practices have been documented since the late 1990s [19], still most systems remain vulnerable. The information systems' complexity, inefficient management practices combined with a lack of trust in the application code stability results in vulnerable systems. Leaving the zero-day exploits aside, most data breaches occur due to unpatched systems of publicly disclosed vulnerabilities [20]. On average, it takes 30-60 days for a vulnerability to be patched that is enough time for adversaries to weaponize and attack targets. For an effective defense, we need to reduce human-in-the-loop interactions in our cyberspace operations, establish trust and reliability in application code.

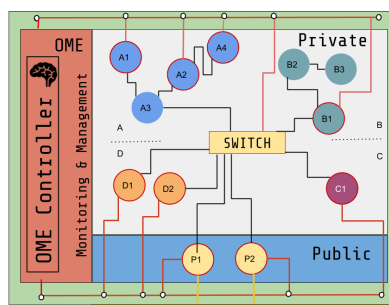


Figure 3
OME Network Model

5.1 Network Operational Model

A network system consists of a set of nodes connected together to form one entity. It provides a set of services using a combination of software running on hardware machines and wetware running on human brains [21]. In the context of an enterprise, wetware can be compared to different job roles. Abstracting network operations with humans, roles, machines, software, and services establish a consistent structure and measurable state for effective automation. To keep our model simple, we will discuss only the machine nodes. The machines can be further categorized into a network or a host node. The network-specific nodes are used for running network operations such as switching, routing, firewall, service proxy or gateway. The host-specific nodes perform actual useful work in the network and provide services such as desktop clients, workstations and application servers. The operational state can be derived locally in a host, for network subnets or for the global network. The local host state provides more context information such as running processes, opened files and user interactions. The network state provides better communication context of the hosts.

Similar to a biological immune system [22], we model a network system with a separate continuous monitoring and control system environment. This Operation Management Environment (OME) provides essential user, data and security operations control. It contains distributed defense machines, agents, installed on nodes that monitor service operations continuously for any control deviations. The OME controller provides centralized command and control for the whole network. In the second-order cybernetics terms, the OME acts as a network system observer. For example, enterprise IT operations are protected using Intrusion Detection & Protection (IDS/IPS) security agents; similar software agents can protect user profile details, provide data privacy and curate content on online social networks services. Figure 3 illustrates this network model. The figure shows the whole network is abstracted into 3 regions - public, private, and OME. The public region exposes services and provides connectivity to the external cyberspace. The private region has internal nodes where services are hosted. The OME region provides monitoring and management for the whole network. Furthermore, the private region can be divided into subregions - A, B, C & D. The OME agents are installed on nodes with red boundaries. Besides capturing local and network operational state, Node B1 also capture network operational state of B2 & B3.

Each node in the network follows - registration, configuration, operation, maintenance, and quarantine life-cycle steps. Furthermore, each OME agent maintains its host operational state and can operate locally on the state without any centralized control. The state can be further divided into configuration, static and dynamic behavior layers. The configuration layer provides secure operational system defaults such as sanitizing default passwords, selecting secure protocol suites, best practices, software hardening rules, and whitelisting to limit unwanted communications. The Configuration Security (ConSec) [23] DARPA program proposes a similar autonomous task. The static layer provides protection from known malware file signatures, malicious network traffic such as

phishing emails and domains. The behavior layer provides runtime knowledge using learning-based methods to model both, normal and abnormal behavior of machine hosts. It is important to note that the static layer expands its signature knowledge as more dynamic threats are confirmed and become common in cyberspace. Figure 4 illustrates these layers with the bottom configuration layer reducing the most attack surface followed by static and dynamic behavior layers.

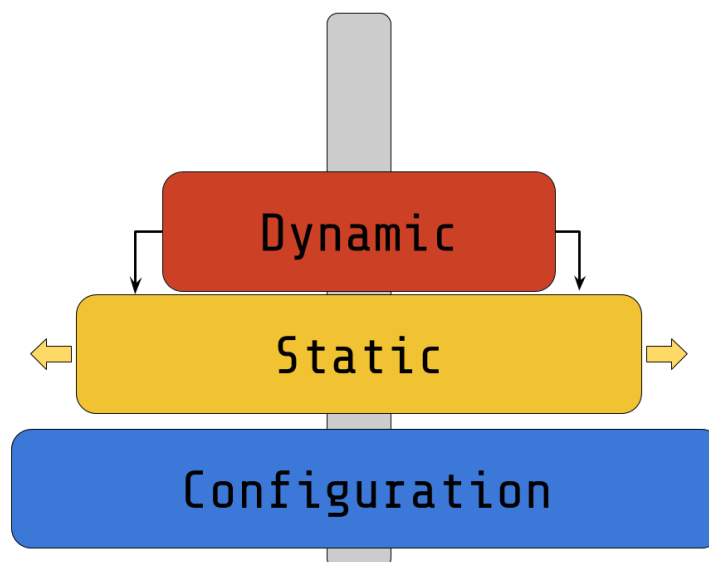


Figure 4
Operational Security Layers

Deception is a critical part of defense operations. It can dupe adversaries to reveal new offensive threats such as zero-day vulnerabilities that a network has not seen before. It helps to increase the static knowledge base. Honeypot networks [24] are isolated real network hosts that provide legitimate services. They replicate real host services and monitor remote user interactions. The OME controller automates the provisioning of honeypot defense agents based upon the real-time operational state. However, AI enabled offensive agents usually scout for honeypot signals before revealing any attack strategies. On any match day, we can watch a cyber deception game between offensive and defensive agents [25].

Sharing emerging threat information using honeypot strategies and malware research amongst OMEs on the Internet can lead to the healthy growth of threat knowledge base and better defenses with early warning signals. For this collaborative information sharing, trusting threat data itself can be an issue, as OMEs are processing data from untrusted network security domains. Effective distributed storage using blockchains [26] can help in creating trust amongst isolated OME security domains. Over time, the threat sharing blockchain network can establish trust based upon a domain interaction history. The OME

creates an active network defense system moving away from current systems that react to threats using protect- monitor-react loops.

The Home Area Network Zero Operations (HANZO) [13] project implements the proposed OME network model. It provides collaborative defense in small network environments having IoT and general-purpose devices by integrating the OME controller in a network gateway. The OME controller maintains the device states and enforces communication endpoints derived from network traffic observations. For instance, a smart light bulb or a power switch may only need to communicate with a local controller, but should not establish any connections to external web services. By learning communication behavior patterns, the OME protects IoT devices from becoming part of a botnet or other malicious activities. The HANZO controller maintains a digital twin for each device in the network to provide automated life-cycle management services of registration, configuration, operation, maintenance, and quarantine. It shares device profiles with other controllers using a blockchain-based global knowledge base. Figure 5 illustrates the internals of the HANZO Controller. The device manager maintains the operational states of network nodes. The agents work in three stages - extraction, generation, and delegation. The extraction stage filters the traffic, the generation stage creates device profiles and continuously monitors for any anomalies. Finally, the delegation stage enforces the device profile in the network.

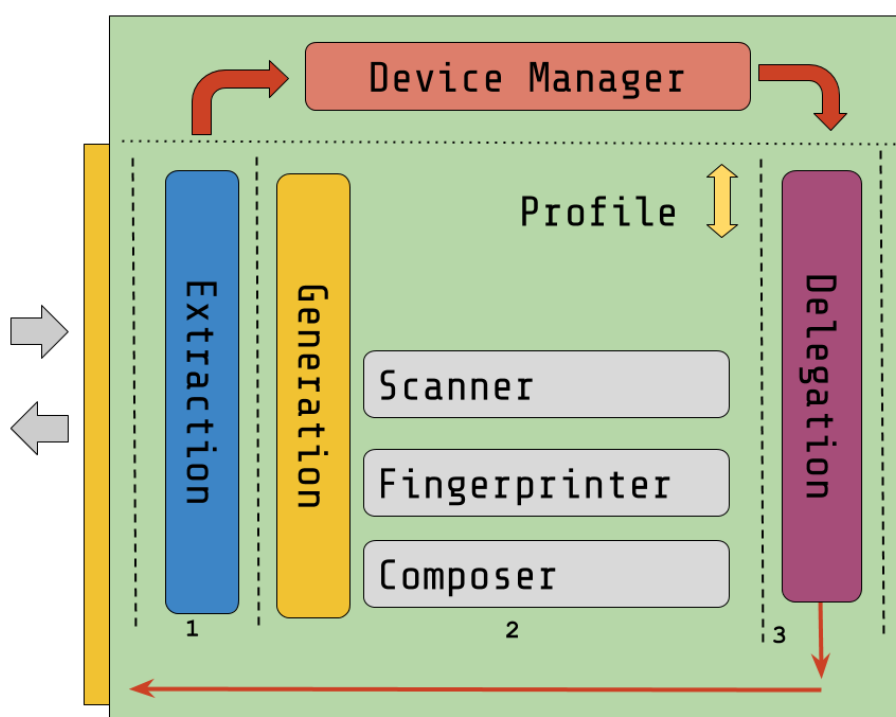


Figure 5
HANZO Controller using OME Model

5.2 Cyberspace Disinformation Defense

Using the OME model, a cyberspace network can be securely monitored and managed. The model assumes information flowing in the network is vetted and represents valid facts about system events. This assumption lacks foundation, and societal-space networks, in particular, political state actors, can exploit this to spread disinformation through media networks resulting in conflicts and information warfare [27].

Citizens consume news media from established news outlets with ethical integrity and spend time in establishing trust in these institutions. With the advent of social media, everyone has become a news source resulting in mass diffusion of fake news. Furthermore, to make users spend more time online, social media networks use personalized information filtering that shows only user preferred content resulting in echo chambers of opinions leaving useful facts behind. The Truth Decay report [28] highlights key issues of disappearing lines between opinion and facts - increased volume of media consumption, diminishing trust in established institutions, and disagreement about facts and analytical data. The report also identifies main actors - (a) academia, research or non-profit organizations that exacerbate research results, (b) media organizations with continuous opinion reporting, (c) foreign nation states to disrupt political systems, and (d) domestic actors to advance personal or political agenda. Detecting disinformation and limiting its spread has become critical for creating a healthy social discourse, if not addressed quickly, we will start to see the foundations of democracy itself becoming insecure.

The current cyberspace media networks are dumb transport pipes lacking any context regarding the veracity of the information they are carrying. The OME model can be applied to online social media networks that can monitor and manage the flow of information. In the model, the environment context has changed with network nodes becoming service users. For example, in a social network service, each user can have an independent instance of OME running, that protects profile data, filter friend posts, and public news items. With information sharing, the OME agents can detect botnets that spread disinformation using flow capacities [29]. For detecting content veracity, agents can perform automated fact checking using NLP [30] and human written facts. Full automated content fact-checking is still an open research problem that will require combined efforts from academia, industry, and government bodies.

5.3 Software Safety and Self-Defense

Everything in cyberspace runs on program code. It is the foundation of all digital life. The hardware and software vulnerabilities can be reduced using secure design, AI-assisted development [31], testing and deployment methods [32]. Trusting a computing stack involves verifying code at each infrastructure layer starting with hardware, bootloader, operating system, device drivers and application software. Automated hardware Trojan implant detection methods [33] can help secure the network hardware layer. Secure-boot chain [34] forms a trust chain by verifying code signature starting from hardware chips, mitigating many attacks like bootkits, rootkits, and persistent malware. Furthermore,

blockchain-based software package management [35] can automate installation and verification of source code.

Using formal mathematical methods to prove software security states establishes trust in the computing base. For example, the seL4 [36] provides a formally verified kernel that can stop attackers from full node compromise. The DARPA deepspec [37] project goal is to create interface specifications and verification of full functional correctness of a software and hardware stack. The OME model, when applied to a software program creates an observer of operation events inside the program context resulting in a self-aware program [38]. We can extend this further to create self-aware host machines. A self-aware system can automatically repair runtime state or repair behavior by writing a patch and updating system code without any downtimes [39]. In 2016, DARPA conducted the Cyber Grand Challenge (CGC) [41] competition to explore similar ideas in automatic defensive systems that are capable of reasoning about flaws, formulating patches and deploying them on a network in real time.

Conclusion

Technology innovations will continue to digitize our society. The autonomous future of societal networks is inevitable. Besides the good uses, every technology has an inherent duality that gives rise to bad uses as well. Using the principles of cybernetics, we propose an autonomous network operations management model that provides real-time defense against emerging threats in cyberspace. The model is shown to be applicable for defenses against disinformation in online social media services. Furthermore, the model can also be applied effectively to protect software code resulting in self-aware system programs. In conclusion, for solving complex security challenges and establishing a peaceful and prosperous society requires all of us to come together. We need ethical programmers, result-oriented academic research, education of citizens, sound and timely policy development by governments, and an industry that puts users first.

References

- [1] Goldstine, H.H. and Goldstine, A., The electronic numerical integrator and computer (ENIAC). Mathematical Tables and Other Aids to Computation, 2(15), pp.97-110, 1946.
- [2] Shannon, C.E., A mathematical theory of communication, Bell system technical journal, 27(3), pp.379-423, 1948.
- [3] Schwab, K., The Fourth Industrial Revolution, World Economic Forum, 2016.
- [4] Society 5.0, Keidanren (Japan Business Federation), 2017.
- [5] Pias, C., The Macy conferences 1946–1953: The complete transactions. Zurich, CH: Diaphanes, 2016.
- [6] Ashby, W. R., An introduction to cybernetics. Chapman & Hall Ltd, 1961.
- [7] Von Foerster, Heinz. "Cybernetics of cybernetics." Understanding understanding. Springer, New York, NY, pp.283-286, 2003.

- [8] The infrastructure of Innovation: Software Defined, Cloud Ready, Intel Corp, 2015.
- [9] Datta, S.P.A., Emergence of Digital Twins, arXiv:1610.06467, 2016.
- [10] Licklider, J.C., Man-computer symbiosis. IRE transactions on human factors in electronics, (1), pp.4-11, 1960.
- [11] Kurzweil, Ray. The singularity is near. Gerald Duckworth & Co, 2010.
- [12] Gunning, David, Explainable artificial intelligence (xai), Defense Advanced Research Projects Agency (DARPA), 2017.
- [13] Singh, Aman, Shashank Murali, Lalka Rieger, Ruoyu Li, Stefan Hommes, Radu State, Gaston Ormazabal, and Henning Schulzrinne. "HANZO: Collaborative Network Defense for Connected Things." In *2018 Principles, Systems and Applications of IP Telecommunications (IPTComm)*, pp. 1-8. IEEE, 2018.
- [14] Antonakakis, Manos, et al. "Understanding the Mirai Botnet." USENIX Security Symposium. 2017.
- [15] Brundage, Miles, et al. "The malicious use of artificial intelligence: Forecasting, prevention, and mitigation." arXiv:1802.07228, 2018.
- [16] Ellison, Robert J., et al. Evaluating and mitigating software supply chain security risks. No. CMU/SEI-2010-TN-016. Carnegie-Mellon Univ Pittsburgh PA Software Engineering Inst, 2010.
- [17] Hutchins, Eric M., Michael J. Cloppert, and Rohan M. Amin. "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains." *Leading Issues in Information Warfare & Security Research* 1.1 (2011): 80.
- [18] Chesney, Robert, and Danielle Keats Citron. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." (2018).
- [19] Anderson, Robert H., et al. Securing the US Defense Information Infrastructure: A Proposed Approach. No. RAND-MR-993-05D/NSA/DARPA. Rand Corp Santa Monica CA, 1999.
- [20] Ablon, Lillian, and Andy Bogart. Zero Days, Thousands of Nights: The Life and Times of Zero-Day Vulnerabilities and Their Exploits. Rand Corporation, 2017.
- [21] Bray, Dennis. Wetware: a computer in every living cell. Yale University Press, 2009.
- [22] Somayaji, Anil. "Immunology, diversity, and homeostasis: The past and future of biologically inspired computer defenses." *Information security technical report* 12, no. 4 (2007): 228-234.
- [23] Torrey, Jacob. "Configuration Security (ConSec)", Defense Advanced Research Projects Agency (DARPA), 2017.
- [24] Libicki, Martin C. "Cyberdeterrence and Cyberwar", Rand Corporation, 2009.

- [25] Jajodia, Sushil, Paulo Shakarian, V. S. Subrahmanian, Vipin Swarup, and Cliff Wang, eds, *Cyber warfare: building the scientific foundation*, Vol. 56. Springer, 2015.
- [26] Ali, Muneeb, Jude Nelson, Ryan Shea, and Michael J. Freedman. "Bootstrapping trust in distributed systems with blockchains." *login: USENIX Mag.* 41, no. 3 (2016).
- [27] Szapranski, Richard. *A Theory of Information Warfare; Preparing for 2020*. AIR UNIV MAXWELL AFB AL, 1995.
- [28] Kavanagh, Jennifer, and Michael D. Rich. "Truth Decay: A Threat to Policymaking and Democracy", RAND Corporation, 2018.
- [29] Shao, Chengcheng, Giovanni Luca Ciampaglia, Onur Varol, Alessandro Flammini, and Filippo Menczer. "The spread of fake news by social bots." *arXiv: 1707.07592* (2017): 96-104.
- [30] Oshikawa, Ray, Jing Qian, and William Yang Wang. "A Survey on Natural Language Processing for Fake News Detection." *arXiv:1811.00770* (2018).
- [31] Allamanis, Miltiadis, Marc Brockschmidt, and Mahmoud Khademi. "Learning to represent programs with graphs." *arXiv:1711.00740* (2017).
- [32] Hummer, Waldemar, Florian Rosenberg, Fábio Oliveira, and Tamar Eilam. "Testing idempotence for infrastructure as code." In *ACM/IFIP/USENIX International Conference on Distributed Systems Platforms and Open Distributed Processing*, pp. 368-388. Springer, Berlin, Heidelberg, 2013.
- [33] Xiao, Kan, Domenic Forte, Yier Jin, Ramesh Karri, Swarup Bhunia, and M. Tehranipoor. "Hardware Trojans: Lessons learned after one decade of research." *ACM Transactions on Design Automation of Electronic Systems (TODAES)* 22, no. 1 (2016): 6.
- [34] Löhr, Hans, Ahmad-Reza Sadeghi, and Marcel Winandy. "Patterns for secure boot and secure storage in computer systems." In *Availability, Reliability, and Security, 2010. ARES'10 International Conference on*, pp. 569-573. IEEE, 2010.
- [35] Nikitin, Kirill, Eleftherios Kokoris-Kogias, Philipp Jovanovic, Nicolas Gailly, Linus Gasser, Ismail Khoffi, Justin Cappos, and Bryan Ford. "CHAINIAC: Proactive software-update transparency via collectively signed skipchains and verified builds." In *26th USENIX Security Symposium (USENIX Security 17)*, pp. 1271-1287. 2017.
- [36] Klein, Gerwin, Kevin Elphinstone, Gernot Heiser, June Andronick, David Cock, Philip Derrin, Dhammika Elkaduwe et al. "seL4: Formal verification of an OS kernel." In *Proceedings of the ACM SIGOPS 22nd symposium on Operating systems principles*, pp. 207-220. ACM, 2009.
- [37] www.deepspec.org , High-Assurance Cyber Military Systems, Defense Advanced Research Projects Agency (DARPA), 2015.
- [38] Hoffmann, Henry, Martina Maggio, Marco D. Santambrogio, Alberto Leva, and Anant Agarwal. "SEEC: a general and extensible framework for self-aware computing." (2011).

[39] Monperrus, Martin. "Automatic software repair: a bibliography." *ACM Computing Surveys (CSUR)* 51, no. 1 (2018): 17.

[40] Andreessen, Marc. "Why software is eating the world." *The Wall Street Journal* 20, 2011.

[41] Frazee, Dustin. <https://www.darpa.mil/program/cyber-grand-challenge>, Defense Advanced Research Projects Agency (DARPA), 2016.

Biography

Aman Singh serves as the Chief Scientist at Palindrome Technologies. He leads the research and development in emerging technologies that have positive societal impacts. His research interests include autonomous information security, fault-tolerant systems, context-aware IoT systems, blockchains, and privacy-aware computing. He holds a graduate degree in Computer Science from Columbia University, New York.

Email: aman.singh@palindrometech.com



Sistemi autonomi simbiotici e sicurezza

Francesco Flammini

Sommario

Il presente articolo indirizza la sicurezza dell'interazione tra macchine ed esseri umani nel contesto dei sistemi autonomi simbiotici (Symbiotic Autonomous Systems, SAS). In particolare, verranno affrontate brevemente le tematiche di sicurezza informatica e dei dispositivi interconnessi, la sicurezza cibernetica ("cybersecurity") e quella dei sistemi cyber-fisici (CPS, Cyber-Physical Systems). Si accennerà infine ai recenti sviluppi della sicurezza in relazione al paradigma dell'Internet delle cose (Internet of Things, IoT).

Abstract

This paper addresses the security of the interactions between machines and humans in the context of Symbiotic Autonomous Systems (SAS). In particular, the paper will provide an overview of the security in interconnected devices, including cybersecurity and Cyber-Physical Systems (CPS) security. We will also mention some recent security developments related to the Internet of Things (IoT).

Keywords: Autonomous Systems, Artificial Intelligence, Internet of Things, Security, Safety, Resilience

Introduzione

Importanza della sicurezza nei SAS

Il nostro ecosistema è molto complesso. Nei sistemi dinamici complessi, piccole modifiche nell'input possono produrre effetti di grandi dimensioni. L'ecosistema può minimizzare molti degli effetti, ma non tutti. Anche i sistemi biologici sono estremamente complessi, in cui i sistemi immunitari mostrano comportamenti molto intelligenti ed adattativi, in grado di combattere la maggior parte degli attacchi di elementi patogeni quali batteri e virus. Tuttavia, accade talvolta che

un microorganismo invisibile possa far oscillare l'omeostasi di un corpo in modo anomalo, fino a debellarlo completamente.

La nostra esistenza non è scontata ed è soggetta ai cambiamenti naturali e indotti dell'ecosistema. Ad esempio, il nostro pianeta sta affrontando una crisi senza precedenti dovuta ai cambiamenti climatici. Sono in fase di avvio massicci programmi di ricerca e sviluppo per realizzare tecnologie e infrastrutture finalizzate alla riduzione delle emissioni di gas serra. Nel celebre incontro del 2009 a Copenaghen, si è concordato di mantenere l'innalzamento della temperatura media al di sotto dei 2°C rispetto al livello preindustriale. Nel 2014, il gruppo intergovernativo sui cambiamenti climatici ha dichiarato che fare ciò richiederebbe una riduzione delle emissioni di gas serra del 40-70% rispetto al livello del 2010 entro il 2050. Si tratta di una sfida molto difficile in termini di impatto sui principali artefici dei cambiamenti climatici: generazione di energia elettrica, trasporti, cibo ed agricoltura [1]. Purtroppo, ognuno dei nuovi sviluppi tecnologici proposti per affrontare tale sfida può potenzialmente essere attaccato e modificato per capovolgere il suo scopo, accelerando così il danno. In passato, era necessario impiantare un agente (di solito una persona) nell'ambiente in modo da realizzare un danno. Oggi l'attacco può essere effettuato tramite Internet ("cyber-attacchi"), e l'Internet of things (IoT) incrementa il rischio di cyber-attacchi quali i tristemente celebri DOS (Denial of Service).

Più specificamente, la scienza della sicurezza informatica affronta i seguenti attributi principali [2]:

- Disponibilità (permanenza del sistema in condizioni di operatività)
- Integrità (dati del sistema non modificati in modo indesiderato)
- Riservatezza (dati riservati non divulgati a persone non autorizzate)

Nel dominio SAS, la disponibilità è estremamente importante poiché i sistemi autonomi simbiotici supportano l'uomo nell'esecuzione di compiti essenziali che potrebbero non essere in grado di effettuare senza l'ausilio di macchine. Ciò è particolarmente rilevante nel campo biomedico (protesi o altri dispositivi di supporto vitale come pompe per insulina intelligenti, pacemaker, ecc.), dove la conseguenza degli attacchi ai SAS può severamente colpire la salute degli esseri umani.

Anche l'integrità è di fondamentale importanza, dal momento che la correttezza dei dati nei SAS può essere molto più critica della loro disponibilità. In molti casi, infatti, risulta preferibile spegnere una macchina, in modalità sicura (es. prevedendo procedure di emergenza manuali), piuttosto che lasciarla funzionare in modo erraneo con conseguenze potenzialmente gravi. Basti immaginare un'automobile o un treno a guida autonoma in presenza di violazioni di integrità: è sicuramente meglio optare per una frenata in sicurezza piuttosto che continuare ad operare in condizioni non sicure.

Infine, la riservatezza è qualcosa di più della semplice "privacy": tenere i dati sensibili fuori dalla portata di utenti non autorizzati e possibili terroristi può impedire il perpetrarsi di minacce quali furti di identità, frodi e altri attacchi

all'integrità e alla disponibilità dei SAS, eseguiti sfruttando le informazioni sottratte.

Nella valutazione del rischio per la "security", tali attributi sono affrontati dal punto di vista di minacce esterne o intenzionali. Tuttavia, guasti sia intenzionali che casuali sono affrontati nel quadro più generale della "computer dependability", tramite le tecniche tradizionali quali ridondanza e tolleranza ai guasti (fault-tolerance), oltre che quelle oggetto di ricerca come l'autoriparazione (self-healing).

A rendere le cose più difficili, i SAS diventeranno sempre più complessi e al tempo stesso critici per la nostra vita ed il nostro benessere [3]. La loro complessità è l'effetto di diversi fattori, tra cui:

- Elevate dimensioni del software necessario per realizzare dispositivi intelligenti ed autonomi, con modalità di guasto non banali e difficili da diagnosticare
- Distribuzione, dovuta alla connettività Internet e alle caratteristiche di mobilità, ubiquità e pervasività (si pensi ai dispositivi indossabili connessi a servizi cloud)
- Eterogeneità dell'hardware, del software, delle connessioni di rete wireless e cablate, dell'impiego congiunto di componenti open-source e COTS (Commercial Off-The-Shelf) insieme ad implementazioni proprietarie e personalizzate.

La criticità deriva dal fatto che i SAS saranno sempre di più impiegati in applicazioni delicatissime per il business, l'ambiente e la salute. Mentre i tradizionali sistemi di sicurezza basati su elaboratori sono mantenuti semplici per facilitarne la verifica ed il collaudo, nonché la certificazione rispetto agli standard internazionali (ad esempio, ISO / IEC 15408 - Common Criteria, IEEE Std 1686-2013 - Standard for Intelligent Electronic Devices Cyber Security Capabilities [4]), la complessità e la limitata prevedibilità dei SAS rappresentano un serio ostacolo. Al tempo stesso, gli ingegneri potranno sfruttare vantaggiosamente l'intelligenza artificiale ed altri paradigmi rilevanti per i SAS (ad esempio, quello dei Digital Twins) per prevenire e combattere le minacce in modo più efficace, basandosi su concetti di proattività, predittività e collaborazione.

Contromisure e loro efficacia

La comunità di ricerca sta facendo un grande lavoro per sviluppare contromisure efficaci per fronteggiare le diverse classi di attacchi informatici. La necessità di competenze nel settore della sicurezza è in costante aumento. Si tratta di competenze fortemente interdisciplinari, che comprendono scienza dei dati, ingegneria elettronica, informatica, analisi del rischio, crittoanalisi, elaborazione dei segnali, intelligence, ecc. È fondamentale, infatti, analizzare non solo le minacce che possono essere generate da determinati individui, ma anche le loro motivazioni, legate ad aspetti umani e socio-culturali che hanno impatto sull'intera comunità.

Tuttavia, poiché il numero di attacchi persistenti aumenta ed è per lo più incontrollato, qualcosa è inadeguato nelle nostre contromisure. In linea di principio, le contromisure basate su osservazioni passate non sono adeguate. Poiché gli attacchi appartengono spesso alla classe dei sistemi dinamici, devono essere combattuti con sistemi dinamici. La ricerca recente in questo settore si basa proprio su tale approccio. In questa breve disanima, passeremo in rassegna alcune tecniche e tecnologie moderne secondo le quali i sistemi autonomi simbiotici dovrebbero essere sviluppati per affrontare le diverse classi di attacchi.

Sicurezza dei sistemi autonomi simbiotici

I SAS sono sistemi dinamici complessi, progettati per esibire comportamenti non banali. Di conseguenza, i SAS sono vulnerabili anche ad attacchi ancora sconosciuti. Proprio come gli esseri umani, si tratta di sistemi complessi, con molti componenti e sottosistemi interagenti e con relazioni non lineari, che operano in un mondo non lineare. Come per gli esseri umani, un semplice virus potrebbe attaccare un SAS e riprodursi rapidamente fino a metterlo fuori uso. Per avere la possibilità di contrastare il virus, il SAS deve possedere una sorta di sistema immunitario in grado di produrre diversi "anticorpi" di test, e quando uno di questi corrisponde all'invasore, il sistema dovrebbe passare alla produzione di un massiccio esercito di "soldati" simili per sopprimere l'invasore in tempo utile.

La sopravvivenza umana è legata alle capacità "intelligenti" che il sistema immunitario mette in atto per sconfiggere gli agenti patogeni. Le difficoltà principali risiedono nella mutazione delle caratteristiche degli agenti patogeni, che li rendono sempre diversi. Proseguendo su tale parallelismo, il sistema immunitario di un SAS può essere ingannato da agenti patogeni mutevoli e deve reagire modificandosi opportunamente. Il problema principale è che i risultati di tali modifiche potrebbero essere imprevedibili e potrebbero portare a fallimenti e comportamenti pericolosi, come quando in un sistema biologico avvengono risposte autoimmuni. Questo è un primo aspetto in cui emerge che l'intelligenza nei sistemi SAS è sicuramente un vantaggio, ma porta con sé aspetti di imprevedibilità che in sistemi critici devono essere attentamente studiati e validati.

Cybersecurity

La cybersecurity è definita come una "disciplina basata sul calcolo che coinvolge tecnologia, persone, informazioni e processi per consentire operazioni sicure. Implica la creazione, la gestione, l'analisi e il collaudo di sistemi informatici sicuri. È un tema interdisciplinare, che comprende aspetti di giurisprudenza, politica, fattori umani, etica e gestione dei rischi con riferimento agli eventi avversi" [5]. La cybersecurity si riferisce al funzionamento sicuro di un sistema informatico, incluse le reti locali, nazionali e globali. Sebbene si tratti di un business da quasi 100 miliardi di euro, ad oggi poche aziende si sentono al sicuro. Ad esempio, nel periodo che va da giugno a dicembre 2009, giganti come Google, Adobe, Yahoo, Symantec, Northrop Grumman, Dow Chemicals, Morgan Stanley hanno subito attacchi provenienti dalla Cina (nome in codice

“Operation Aurora”). A tal proposito, il gigante dell'informatica Google ha introdotto la cosiddetta politica ZeroTrust: nessuna persona esterna o interna all'organizzazione è considerata completamente affidabile.

Non esiste una serie standard di regole utilizzate per affrontare il crescente numero di minacce provenienti da hacker, ransomware e furto di dati. Negli Stati Uniti, il National Institute of Standards and Technology (NIST) ha pubblicato il Cybersecurity Framework [6] ed una roadmap per migliorare la cybersecurity delle infrastrutture critiche. Il framework identifica cinque funzioni per organizzare un sistema di sicurezza: identificare, proteggere, rilevare, rispondere e ripristinare. Il framework è destinato prevalentemente alle piccole imprese con lo scopo di deframmentare il mondo della sicurezza informatica. Il dipartimento americano della sicurezza nazionale (Homeland Security) identifica 16 settori infrastrutturali critici tra cui reti di telecomunicazioni, difesa, protezione civile, servizi finanziari, alimentazione, sanità, reattori nucleari, sistemi di trasporto e reti idriche [7].

La principale motivazione degli attacchi informatici risiede nelle vulnerabilità del software che possono essere sfruttate dagli hacker. Gli incidenti di cybersecurity più comuni sono dovuti a malware e virus convenzionali (64%), attacchi di ransomware (30%), errori dei dipendenti e azioni involontarie (27%) [8]. Per farsi un'idea dell'entità del problema, basti citare che circa il 40% dei computer con connessione Internet condivisa presenta attacchi ogni sei mesi, la maggior parte delle organizzazioni (61%) hanno una gestione della sicurezza inadeguata, ed il 66% delle imprese manifatturiere non ha budget dedicato alla sicurezza informatica.

Sicurezza cyber-fisica

La sicurezza cosiddetta cyber-fisica si riferisce alla sicurezza di sistemi informatici operanti in stretta connessione con un ambiente fisico, come i sistemi di controllo industriale (Industrial Control Systems, ICS), le auto a guida autonoma ed i droni, in cui i componenti cyber hanno responsabilità di rilevamento, controllo e azionamento, e sono connessi a Internet. A causa della connessione Internet, il sistema diventa vulnerabile alle minacce alla sicurezza che possono verificarsi nelle "reti aperte". La differenza tra la cyber-security e la cyber-physical security è che mentre nella prima le conseguenze si esauriscono il più delle volte nel modo cyber, con eventuali ripercussioni indotte in modo indiretto nel modo fisico, nella seconda le conseguenze sono direttamente impattanti sul mondo fisico, anche se la parte cyber risultasse apparentemente non compromessa. In altre parole, il danno potrebbe limitarsi al mondo fisico, con conseguenti difficoltà in termini di diagnosi e rilevamento nella parte cyber. Alcuni esempi riportati nel seguito chiariranno meglio questo concetto. In entrambi i casi, però, è importante chiarire che la “sicurezza fisica”, ovvero la protezione nei confronti di accessi e manipolazioni del sistema di tipo tangibile (es. tramite accesso non autorizzato ai locali tecnici del datacenter di un impianto industriale) rimane un aspetto distinto, per quanto fortemente correlato al concretizzarsi di talune minacce: si

pensi ad es. a connessioni cablate su reti isolate, in cui l'accesso fisico alle reti e ai dispositivi è indispensabile.

Dal momento che i sistemi autonomi simbiotici incorporano dispositivi fisici connessi per interagire con gli esseri umani e l'ambiente, essi appartengono anche alla categoria dei sistemi cyber-fisici (CPS, Cyber-Physical Systems) e sono pertanto soggetti alla valutazione del rischio tipica di questi sistemi.

Rischi, vulnerabilità e conseguenze nei sistemi cyber-fisici

Essendo dotati di componenti sia fisici che cyber, i CPS possono essere soggetti a minacce, vulnerabilità e conseguenze che investono entrambi i mondi.

Un esempio di attacco ad un CPS è quello di una centrifuga la cui operatività può essere modificata causando un rallentamento del processo quasi impercettibile. È quello che è successo con l'attacco del virus "Stuxnet" alle centrifughe iraniane di un impianto nucleare controllate da un sistema SCADA (Supervisory Control and Data Acquisition), che ne causò danni ingenti ed infine l'arresto nel 2010 [9]. Un sistema SCADA è un'applicazione che consente agli operatori umani di monitorare un processo industriale e di archiviare e analizzare i dati di processo corrispondenti. Si tratta di un sistema critico in quanto attacchi di tipo cyber, come quello menzionato, possono avere conseguenze devastanti sui componenti fisici e finanche causare disastri ambientali.

Un altro esempio è quello della cosiddetta "smart-grid" (rete intelligente di distribuzione dell'energia elettrica), in cui è necessario monitorare le utenze al fine di ottenere dati in tempo reale relativi per ottimizzare l'erogazione di energia verso specifiche destinazioni. Contatori automatici, intelligenti ed interconnessi (smart-meters), sono utilizzati per misurare il consumo energetico dei consumatori e per fornire una varietà di servizi a valore aggiunto. Poiché le informazioni raccolte possono includere dati sensibili dei consumatori, queste devono essere protette opportunamente, il che rappresenta una criticità di tipo "privacy" comune a molti sistemi cyber-fisici (si pensi ad esempio a quelli impiegati in applicazioni "smart-health" di monitoraggio remoto dei pazienti), e che deve pertanto essere annoverata tra gli aspetti a cui dedicare particolare attenzione nelle analisi di sicurezza di tali sistemi [10].

Altri esempi rilevanti di sicurezza di sistemi CPS sono:

- L'alterazione dell'operazione e l'esplosione di un impianto petrolchimico (evento realmente avvenuto in Arabia Saudita nel 2017).
- Il controllo remoto di un'automobile connessa di ultima generazione, dotata di pilota automatico, con conseguenze potenzialmente letali.
- La modifica del software di una macchina per la risonanza magnetica connessa al fine di generare immagini errate del cervello.
- La modifica di una pompa per infusione collegata a Internet al fine di somministrare una quantità letale di medicinale al paziente.
- La manipolazione di un frigorifero connesso al fine di deteriorare gli alimenti e causare intossicazioni ed infezioni.

- La manipolazione del sistema di controllo di una diga con conseguenze devastanti per milioni di persone (ci sono circa 100.000 dighe solo negli Stati Uniti).
- Lo spegnimento dell'impianto elettrico in una casa, in un distretto o nell'intera città con le conseguenze che è facile immaginare.

Poiché la maggior parte degli attacchi di cui sopra sono stati provati ed alcuni di essi hanno avuto successo, risulta evidente che è necessario incrementare il livello di sofisticazione delle contromisure. È altresì evidente che contromisure puramente reattive non sono sufficienti: i SAS, in particolare, dovrebbero essere sviluppati per affrontare tali attacchi cyber-fisici in modo pro-attivo, predittivo e collaborativo. In altre parole, i futuri CPS dovrebbero essere progettati non solo per essere sicuri ed affidabili nell'affrontare minacce note, ma anche per riconoscere minacce sconosciute, adattarsi ed organizzarsi per fronteggiarle in modo possibilmente strategico e ripristinare il normale funzionamento nel più breve tempo possibile. Ci si riferisce spesso a queste ultime caratteristiche come "self-healing" e "resilienza".

Problemi legati al supporto e all'eterogeneità dei fornitori

I produttori di software forniscono supporto per un prodotto per un ragionevole lasso di tempo, mentre le aziende di hardware forniscono supporto per un periodo di tempo molto più breve (2-3 anni), con la situazione che è iniziata a cambiare nel 2018 (7-10 anni). Si tratta di un problema importante, in quanto i prodotti connessi possono generare piccoli profitti richiedendo anni di aggiornamenti e patch per adeguarne la sicurezza.

Inoltre, molti dispositivi connessi sono assemblati con componenti (hardware, firmware, software) sviluppati da diversi fornitori. Ciò è particolarmente preoccupante perché se il collegamento più debole non viene aggiornato regolarmente, l'intero sistema diventa vulnerabile. Mentre nella maggior parte dei sistemi IT omogenei vi è un processo di aggiornamento regolare del software interamente sotto il controllo di un fornitore, nei sistemi cyber-fisici eterogenei ciò non è ancora regolato in modo chiaro e coordinato. Ciò si riflette inevitabilmente sui sistemi autonomi simbiotici che incorporano dispositivi fisici interconnessi vulnerabili ad attacchi cyber-fisici. È pertanto necessario approfondire energie sullo sviluppo di standard, metodologie e strumenti che possano supportare il patching e l'aggiornamento regolare dei SAS eterogenei.

Impatto della sicurezza cyber sulla società

Per tutto quanto scritto precedentemente, è evidente che la cybersecurity può avere un notevole impatto economico e sociale su intere comunità. Le conseguenze di attacchi possono essere ancora più rilevanti e riguardare salute ed ambiente laddove la sicurezza riguarda i sistemi cyber-fisici. Ma c'è di più.

Nella società moderna, gran parte delle interazioni sociali sono veicolate tramite Internet. Si pensi ad esempio all'uso massiccio di motori di ricerca, blog, e social networks. Molti dei contenuti che ci raggiungono, le notizie e gran parte delle risposte alle nostre domande quotidiane sono pertanto "filtrati"

da entità cyber secondo algoritmi a noi sconosciuti e sotto il controllo di altri. Ad esempio, Flipboard e applicazioni simili selezionano gli articoli in base ai nostri interessi. Amazon e sistemi simili di commercio elettronico (e-commerce) suggeriscono acquisti di libri, musica, video e prodotti fisici in base alle nostre precedenti interazioni e ai profili social. Nella società futura, la percezione della realtà sarà potenziata ("realtà aumentata") attraverso dispositivi portatili ed indossabili (si pensi agli smart-glasses) che fungeranno sempre più da vere e proprie "protesi". Già oggi il fenomeno delle "fake news" ed altri condizionamenti indotti sono oggetto di discussione a tutti i livelli, ed è quindi chiaro che tali manipolazioni potranno essere usate sempre più per distorcere la comprensione della realtà, in modo da sviluppare una realtà alternativa secondo la strategia dell'attaccante. L'alternanza di informazioni affidabili con risposte fasulle rappresenta un ulteriore elemento che può sviare la comprensione della realtà. Con un opportuno innesco della realtà distorta, le notizie false possono sembrare plausibili e superare l'ostacolo della capacità critica dell'individuo, tra l'altro destinata tanto più ad atrofizzarsi quanto più le tecnologie sostituiranno i suoi strumenti percettivi.

Nel campo dell'istruzione, il processo di distorsione delle informazioni ha qualche possibilità di auto-correzione per via delle capacità di discernimento degli studenti. Se invece la distorsione viene applicata ad informazioni quali cartelle cliniche, rapporti di produzione, relazioni finanziarie oppure posizioni politiche impiegate nel processo decisionale, le conseguenze potrebbero essere molto più rilevanti. In situazioni basate sulla fiducia, in cui non vi è tempo per la verifica di qualità delle informazioni fornite, l'alterazione potrebbe essere decisiva.

Standard e sicurezza nell'Internet of Things

Questa sezione descrive in che modo l'industria sta rispondendo ai rapidi sviluppi nell'area della connettività della Internet of Things (IoT) e di alcuni sistemi di protezione dei dati.

Il collegamento dei dispositivi IoT ad una rete locale o ad Internet richiede tecnologie cablate e wireless sviluppate molto prima che il concetto IoT fosse introdotto e non sempre adeguate a supportarne i requisiti. Le soluzioni wireless possono comprendere protocolli proprietari a livello di dispositivo, connessioni WiFi, Bluetooth Low Energy (BLE), rete cellulare, ed altre tecnologie. La connettività del dispositivo e molte altre elaborazioni di dati appartengono a ciò che viene definito "edge computing" (livello intra-dispositivo) o "fog computing" (livello area locale o metropolitana), in opposizione al "cloud computing" che prevede l'utilizzo di server in data center remoti raggiungibili tramite Internet, la cui localizzazione non è sempre ben definita. È intuitivamente evidente che il numero teorico di potenziali minacce sale man mano che ci si sposta dal livello locale a quello geografico, su cui si ha meno controllo a livello di infrastruttura, ed è altresì proporzionale al numero di collegamenti wireless impiegati, per via della maggior "apertura" e accessibilità fisica di tali reti. Vi sono però validi controesempi: un provider serio di servizi cloud può fornire maggiori garanzie di sicurezza rispetto ad un

servizio locale mal gestito; un'infrastruttura di rete locale cablata in cui i dati scambiati non sono crittografati (per via di una eccessiva confidenza sul livello di segregazione fisica) può risultare all'atto pratico meno sicura di una rete WiFi dotata di adeguata cifratura; per motivi analoghi, il completo isolamento da Internet può essere un'arma a doppio taglio in quanto impedisce il regolare aggiornamento dei dispositivi con le più recenti patch di sicurezza.

IoT LPWAN

Il LoRaWAN Alliance ha annunciato l'implementazione di un protocollo di rete geografica a bassa potenza (LPWAN) che permette il collegamento wireless a Internet di oggetti alimentati a batterie, a livello di reti regionali, nazionali, o globali. Dispone di comunicazioni dati bidirezionali e di sicurezza end-to-end, servizi di mobilità e localizzazione. Utilizza la topologia "star-of-star" in cui i gateway inoltrano i messaggi tra i dispositivi finali ed il server di rete centrale. Il gateway funge da ponte transponder per convertire i pacchetti a radiofrequenza (RF) nei pacchetti IP (Internet Protocol). È dotata di "Firmware-Over-The-Air" (FOTA) per aggiornare i dispositivi collegati o per distribuire i messaggi. Esempi di dispositivi sicuri LoRa comprendono Cypress PSoC 6MCU, Semtech, e ARM Cortex-M 2MCU.

WiFi IoT (IEEE 802.11)

Se il fattore consumo energetico non è critico, la connettività IoT può utilizzare un'infrastruttura ben consolidata di hub e router WiFi. Ad esempio, nel 2018, NXP Semiconductor ha sviluppato un dispositivo edge-computing IoT-on-a-Chip basato sul processore applicativo ARM iMX 6ULL, completo di WiFi (dual-band 802.11ac) e Bluetooth (4.2). Ha un ingombro ridotto (14x14x2,4mm), è scalabile e facile da usare nei progetti. Sarà seguito da iMX7 e iMX8 nel 2019. L'IoT-on-a-Chip è adatto per l'impiego in robot sicuri, nel rilevamento e gestione manomissioni, e nella crittografia ad alto throughput. Può anche essere espanso con un'interfaccia inter-chip personalizzata.

IoT ZigBee (IEEE 802.15.4)

Per le applicazioni che non esigono trasferimenti dati ad alta velocità, ma richiedono una bassa potenza, i dispositivi di rete auto-organizzanti per ambienti industriali utilizzano spesso il protocollo ZigBee. Ad esempio, la STMicroelectronics ha annunciato un chip wireless a doppio processore, STM32WB System-on-Chip (SoC) per supportare il protocollo ZigBee. È basato sul microcontrollore ARM Cortex-M4 per l'esecuzione dell'applicazione principale. Un altro core ARM Cortex-M0+ viene utilizzato per alleggerire il processore principale. Prevede il Bluetooth Low Energy (BLE) 5 ed una radio ZigBee in grado di eseguire non solo il ZigBee ma anche l'OpenThread e altri protocolli. Un'altra particolarità è il "balun" interno incluso sul SoC per il collegamento diretto all'antenna con risparmio di 6 elementi esterni. La radio a bassa potenza a 2,4 GHz richiede 5,5 / 3,8 mA per trasmettere / ricevere.

IoT Bluetooth Low Energy (BLE) (IEEE 802.15.4)

Le applicazioni in ambito sanitario, fitness, sicurezza e home entertainment richiedono spesso connettività edge IoT a basso consumo attraverso la rete BLE

“personal area network” (PAN). Il BLE va bene per dispositivi che devono rimanere operativi per un anno prima della successiva ricarica. La copertura è regolabile da 3m a 60m circa. Ad esempio, Texas Instruments ha annunciato numerosi dispositivi SimpleLink BLE. Tra questi, lo MCU CC2642R è adatto ad applicazioni Bluetooth 4 e 5 in dispositivi 2.4 GHz e Sub-1 GHz, con un assorbimento in modalità “sleep” inferiore a 1µA. Ha una CPU ARM Cortex-M4F a 48 MHz, un ARM Cortex-M0 dedicato per il controllo radio, oltre ad un Sensor Controller a bassissimo assorbimento per sensori digitali e analogici, ed acquisizione dati. Il suo ricetrasmittitore può gestire molti protocolli compresi WiFi, BLE, ZigBee, Thread, Sub-1-GHz ed Ethernet.

IoT Cellular Networking

Molti dispositivi IoT richiedono connettività cellulare, tramite GSM, CDMA, LTE in ambienti 2G, 3G, 4G e 5G. Tra le aziende che operano nel settore, Nortfic Semiconductor ha annunciato la sua nRF91 per lavorare con Verizon Wireless Network (USA) e Telia (Norvegia). Il nRF91 è un System-in-Package (SiP) multimodale LTE-M / NB-IoT ultracompatto, a bassa potenza e piccolo ingombro (10x161,2mm). Utilizza il processore ARM Cortex-M33, il processore di sicurezza ARM TrustZone e il GPS assistito. Ha un modem, un ricetrasmittitore SAW-less e un frontend Qorro RF. Il sistema nRF91 è destinato non solo agli attuali smartphone, ma anche a molti altri dispositivi mobili. Esso comprende funzioni di sicurezza per l'hardware ed il software applicativo attraverso l'ARM Cortex-M33 e l'ARM CryptoCell-310. Il TrustZone per ARMv8-M protegge i dati, il firmware e le periferiche.

IoT Near-Field Communications (NFC)

La tecnologia di comunicazione cosiddetta “near field” (NFC) consente di collegare un dispositivo portatile con un terminale in modalità contactless. Viene stabilita una connessione quando il dispositivo portatile viene avvicinato al terminale senza necessità di contatto fisico. Esempi di protocolli utilizzati sono Bluetooth ed il FeliCa. Gli standard utilizzati sono ISO / IEC 14443 (106 kb/s) per le carte d'identità e ISO / IEC 18000-3 per i dispositivi RFID. A titolo di esempio, la STMicroelectronics ha annunciato i tag ST25NF come IC Tag Type 4 (ST25TA), Dynamic Tag IC tipo 5 (ST25DV) e IC Tag 5 (ST25TV), ora certificati.

Oltre a quelli elencati, esistono diverse altre implementazioni di connettività a livello “edge”, e tutti prevedono approcci simili di “security by design” (progettazione basata sui requisiti di sicurezza).

Esempi di ricerca in sicurezza

Poiché gli attacchi informatici sono molto sofisticati, non sono applicabili tecniche semplici per il loro rilevamento. Si stanno conducendo ricerche approfondite in tutto il mondo per sviluppare tecniche efficaci capaci di rilevare anomalie su reti connesse attraverso l'analisi dei pacchetti trasmessi, ed altri strumenti finalizzati ad implementare contromisure idonee. Quasi tutti i dati catturati da una rete sotto attacco possono considerarsi non stazionari. Un'analisi approfondita è necessaria per segmentare i segnali in frame adeguati, estrarne le caratteristiche (“feature”) più significative,

assemblare le feature in vettori, inserire i vettori in un classificatore, classificare il comportamento della rete in base alle feature, ed infine determinare contromisure appropriate per proteggere la rete ed i dispositivi ad essa connessi. L'acquisizione e la segmentazione dei dati è il prerequisito per l'analisi delle feature. L'estrazione di feature dai dati costituisce la prima grande sfida e viene spesso eseguita utilizzando analisi temporali standard e/o spettrali mono-scala. C'è fermento nella ricerca finalizzata a migliorare l'estrazione delle feature attraverso tecniche di elaborazione del segnale multi-scala e poli-scala [11]. Oltre alle metriche tradizionali basate sull'energia, si utilizzano anche metriche basate sull'entropia in grado di estrarre modelli ("pattern") relativi alle informazioni presenti nei dati [12] [13]. Le metriche devono anche essere adatte per il traffico di tipo burst ("a raffica"), che è caratterizzato da una dipendenza a largo spettro, con distribuzioni di probabilità-massa-funzione dotate di lunghe "code".

Il secondo elemento principale nello sviluppo della tecnica è identificare l'inizio dell'attacco. Il terzo è classificare gli attacchi e lanciare contromisure appropriate. Le fasi di rilevamento e classificazione coinvolgono tecniche di apprendimento automatico ("machine learning") e apprendimento profondo ("deep learning") [14]. Recentemente, sono state condotte indagini sui tali metodi per l'analisi delle reti al fine di rilevare intrusioni, corredate dalla descrizione dei dataset di rete comunemente impiegati [15]. Esempi di tecniche di rilevamento anomalie per contrastare i famosi attacchi DDoS (Distributed Denial of Service) sono forniti nei riferimenti bibliografici [16] e [17].

Sfide e problemi aperti per la sicurezza dei SAS

I SAS sono sistemi basati su computer appartenenti alle classi dei sistemi embedded ed in particolare dei CPS cosiddetti "smart", ovvero dotati di intelligenza avanzata per renderli autonomi ed in grado di operare simbioticamente tra loro, con gli esseri umani e con l'ambiente. In quanto tali, la loro sicurezza eredita la maggior parte delle metriche, delle vulnerabilità emergenti e delle tecnologie di protezione sviluppate per le classi di sistemi a cui appartengono. Sebbene la complessità e la criticità dei SAS possano essere superiori ai più comuni sistemi di elaborazione, il che complica notevolmente la loro progettazione ed analisi, essi presentano alcune funzionalità uniche che difficilmente possono essere ritrovate nei sistemi tradizionali. Tali capacità sono una conseguenza della loro maggiore intelligenza e autonomia, che possono essere vantaggiosamente sfruttate per progettare macchine che si auto-diagnosticano, auto-proteggono e auto-riparano ("self-healing"). Inoltre, essi cominceranno presto a mostrare alcuni aspetti di sicurezza proattiva: invece di assumere che ogni cosa nell'ambiente circostante stia lavorando secondo specifica, essi inizieranno a prevedere i comportamenti anomali nell'ambiente da parte dei loro simili e degli esseri umani [18].

Ad esempio, si consideri il caso di veicoli autonomi dotati di infrastruttura di controllo centralizzata che si avvicinano ad un incrocio. Secondo i tradizionali paradigmi di sicurezza, il sistema di controllo centralizzato consentirà solo ad

alcuni veicoli di muoversi e comanderà agli altri di rimanere fermi. I veicoli autorizzati a muoversi ignoreranno completamente la presenza degli altri veicoli (come nel caso del controllo ferroviario) o al massimo individueranno ostacoli sul loro percorso se dispongono di sensori appropriati o di visione artificiale. Pertanto, negli attuali sistemi critici per il controllo del traffico, comportamenti anomali dei veicoli che non sono previsti nelle specifiche tecniche possono facilmente portare ad incidenti. Immaginiamo ora cosa succede quando noi esseri umani ci avviciniamo ad un incrocio e osserviamo che un veicolo si sta muovendo ad alta velocità lungo una strada laterale verso la nostra traiettoria. Anche sapendo che quel veicolo sarà obbligato ad arrestarsi se abbiamo noi la precedenza, rallenteremo prudentemente in quanto percepiamo un livello di rischio più elevato, associato al fatto che quel veicolo potrebbe non essere in grado di fermarsi. Lo stesso vale per qualsiasi altro comportamento esterno che percepiamo come non sicuro o minaccioso. Ciò non avviene nella maggior parte degli attuali sistemi di controllo computerizzati, a meno che non vi sia un requisito specifico, dal momento che le regole funzionali sono normalmente molto semplici e facili da valutare. Tuttavia, è quasi impossibile prevedere tutto ciò che può accadere in scenari reali, e gli incidenti più gravi possono accadere come conseguenza di situazioni imprevedibili o impreviste; ad esempio, a causa di hacker che assumono il controllo di alcune entità di sistema. La capacità dei SAS di adattarsi e reagire a scenari imprevisti è pertanto una caratteristica che può controbilanciare le debolezze legate alla loro stessa imprevedibilità, portando ad un apparente paradosso. Ciò significa che è necessario un cambio di paradigma per passare dal tradizionale controllo centralizzato basato su sicurezza predicibile, ai sistemi autonomi decentralizzati ed adattativi, in grado di reagire a minacce sconosciute e scenari operativi imprevisti.

Analogamente, e in modo diverso dagli esseri umani, i sistemi attuali non sono progettati per, e quindi incapaci di, fare ciò che sarebbe altrimenti desiderabile, come ad esempio:

- Rilevare comportamenti anormali in agenti paritari, ed avvisare prontamente gli altri agenti con cui collaborano.
- Avvisare i loro simili sui pericoli e le minacce che potrebbero incontrare, in base alla loro esperienza e conoscenza accumulata.
- Supportare le altre entità nel difendersi e riprendersi dagli attacchi.
- Cooperare strategicamente e coordinare gli altri per rispondere meglio agli attacchi.

Mentre un simile livello di intelligenza e autonomia può apparire fantascientifico, molto è già stato fatto per raggiungere un livello almeno moderato di sicurezza proattiva, facendo leva sull'apprendimento automatico, sul riconoscimento delle situazioni ("situation awareness") e sul supporto decisionale. Ad esempio, il rilevamento delle intrusioni nelle reti utilizza approcci euristici per rilevare anomalie nel traffico di rete e nel comportamento degli utenti. Modelli di riconoscimento basati su reti bayesiane o su reti neurali

artificiali possono essere addestrati al fine di mettere a punto le loro prestazioni [19].

A un livello cognitivo più avanzato, il concetto di "distributed reflective architectures" è stato introdotto per la prima volta circa 15 anni fa, ed è stato in qualche modo pionieristico rispetto all'attuale ricerca sui gemelli digitali ("digital twins") [20]. In tali modelli concettuali, gli agenti autonomi presentano una qualche forma di monitoraggio e controllo reciproco. Affinché un SAS possa monitorare i suoi simili e l'ambiente, è necessario che adeguati modelli predittivi siano inclusi nel suo "cervello". Limiti di capacità e prestazioni hanno impedito di gestire tali modelli computazionali complessi in passato, ma in futuro opportune combinazioni di edge, fog e cloud computing consentiranno a dispositivi connessi con limitate potenze di calcolo di gestire le necessarie elaborazioni in modo efficiente, con il miglior compromesso tra precisione, tempi di risposta e consumo energetico. I SAS prevederanno dei cloni virtuali di loro stessi, delle altre entità con cui interagiscono e dell'ambiente, al fine di eseguire simulazioni accelerate "what if?" basate sulla verosimiglianza di minacce, attacchi e altri eventi indesiderati. Proprio come fanno gli umani quando si preoccupano delle implicazioni di un potenziale pericolo, i SAS valuteranno le situazioni, avvertiranno gli umani, prediranno le possibili conseguenze e pianificheranno rapidamente contromisure efficaci. Inoltre, livelli più alti di intelligenza consentiranno ai SAS di impostare autonomamente strategie di sicurezza avanzate, come trappole e mimetizzazioni, al fine di fuorviare gli avversari.

Conclusioni

La cybersecurity, la sicurezza cyber-fisica, e le relative implicazioni sociali sono sotto la lente d'ingrandimento del mondo accademico, della ricerca industriale, dei governi, e di tutti i cittadini preoccupati di quello che sarà il futuro dei propri figli. Si tratta di una delle sfide più grandi e avvincenti, che riguarda le fabbriche, la finanza, la sanità e le maggiori organizzazioni, e che ha un impatto su tutti gli aspetti della nostra esistenza, da quelli tangibili (es. salute, protezione) a quelli cognitivi (es. percezione, opinioni). Sono stati pubblicati migliaia di articoli tecnici, libri e riviste che trattano argomenti correlati alla sicurezza dei sistemi SAS, presenti negli archivi dei più noti editori (IEEE/Wiley, ACM, Springer, Elsevier, ecc.), ma nonostante ciò quello della sicurezza dei sistemi intelligenti autonomi è a tutt'oggi un campo di ricerca apertissimo con molte questioni irrisolte, non solo di tipo tecnico, ma anche etico, sociale e legale. Si pensi all'accettabilità sociale dei "cobots" (robot collaborativi che affiancano gli umani nelle catene di montaggio) potenzialmente pericolosi, alle responsabilità legali degli incidenti causati da sistemi autonomi imprevedibili (oltre che "indottrinabili"), ed infine ai "dilemmi" che si presentano quando un veicolo stradale a guida autonoma, in una situazione critica, dovrà necessariamente scegliere se sacrificare i pedoni o l'equipaggio del mezzo.

Evidentemente, quello della cybersecurity è un ambito multidisciplinare, richiede una solida base in ingegneria ed informatica, ma non solo. Molte università hanno preso in considerazione l'idea di intraprendere programmi

specifici di cybersecurity. Tuttavia, le conoscenze fondamentali su cui si sta sviluppando tale campo sono frammentate, e pertanto stanno prendendo piede iniziative come quella del progetto CyBOK (Cyber Security Body of Knowledge) allo scopo di stabilire le conoscenze fondamentali e generalmente riconosciute sulla sicurezza informatica [21]. Un'iniziativa notevole in quest'ambito, che ne testimonia l'importanza e l'attualità, è quella dell'IEEE che con il suo Future Directions Committee ha istituito una specifica "IEEE Cybersecurity Initiative", in cui è stato sviluppato il progetto Try-CybSi [22]. Inoltre, nel 2015 è stata istituita una task force congiunta sull'educazione alla cybersecurity ("Joint Task Force on Cyber-Security Education"). Tra le varie iniziative connesse delle organizzazioni di settore, si menziona il comitato tecnico IFIP (International Federation for Information Processing) sull'Information Security Education (IFIP WG 11.8). Nonostante i molti sforzi profusi per specificare un quadro educativo omogeneo per la sicurezza, il materiale didattico differisce sostanzialmente tra le varie istituzioni. Anche i requisiti di formazione dell'industria variano notevolmente e pertanto alcune organizzazioni tecniche hanno formulato delle linee guida a riguardo (vedi ad es. [23]).

Tutti questi sforzi, tuttavia, dovrebbero essere accompagnati dal cambio di paradigma abilitato dalla futura generazione di SAS. Infatti, nel controllo distribuito attuale, le entità sono prevalentemente progettate per:

- Fidarsi incondizionatamente dei propri supervisori
- Non preoccuparsi dei loro simili e dell'ambiente circostante, fatta eccezione per le variabili che sono tenute a monitorare
- Comandare e dimenticare ("set and forget") le entità da loro supervisionate

Al fine di raggiungere una resilienza a 360 gradi, oltre alle funzionalità di autodiagnostica e autoriparazione, i SAS saranno chiamati a:

- Controllare criticamente se possono fidarsi del loro supervisore, proprio come fanno gli umani, verificando se i comandi che ricevono siano ragionevoli e sicuri
- Porre attenzione ai comportamenti dei loro simili e alle evoluzioni dell'ambiente circostante, al fine di rilevare eventuali anomalie e comportamenti anormali che potrebbero essere sintomo di sabotaggio
- Verificare in modo proattivo se le entità supervisionate stiano effettivamente eseguendo le attività che sono state loro assegnate, o se qualcun altro ha preso il loro controllo

Ciò non richiede necessariamente che i SAS siano in grado di riconoscere e classificare situazioni sconosciute. Un modo fattibile per ottenere tali risultati è quello di incorporare modelli (semplificati) di altri SAS interagenti, imitando il comportamento simile degli esseri umani e cominciando ad impiegare in modo strutturato il paradigma dei "digital twin". Una cooperazione intelligente per riconoscere meglio le minacce e rispondere rapidamente ai malfunzionamenti è uno dei principali risultati altamente auspicabili nella ricerca sui SAS sicuri [24] [25].

Riferimenti

1. IEEE, Spectrum Magazine, June 2018. Available Aug 15, 2018 from IEEE at <https://spectrum.ieee.org/magazine/2018/June>
2. Algirdas Avizienis, Jean-Claude Laprie, Brian Randell, and Carl Landwehr, "Basic concepts and taxonomy of dependable and secure computing," IEEE Trans. Dependable and Secure Computing, vol. 1, no. 1, pp. 11-33, Jan-Mar 2004. (doi: 10.1109/TDSC.2004.2).
3. Francesco Flammini (ed.), Resilience of cyber-physical systems: From risk modeling to threat counteraction. Series: Advanced Sciences and Technologies For Security Applications. Berlin-Heidelberg: Springer, 2019 (ISBN 978-3-319-95597-1, eBook; 978-3-319-95596-4).
4. International Standards Organization (2009). ISO/IEC 15408-1:2009 - Information technology -- Security techniques -- Evaluation criteria for IT security -- Part 1: Introduction and general model. [online] Iso.org. Available at: <https://www.iso.org/standard/50341.html> [Accessed 24 Sep. 2018]
5. JTFSCSE, Joint Task Force on Cybersecurity Education (CSE). 2015. Available Aug 15, 2018 from IEEE at <https://www.csec2017.org/>
6. NIST, Cybersecurity Framework. Version 1.1. Gaithersburg, MD: The National Institute of Standards and Technology (NIST), Apr 16, 2018. <https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf>
7. DHS, Critical Infrastructure Sectors. Homeland Security. Aug 14, 2018. Available Aug 15, 2018 from IEEE at <https://www.dhs.gov/critical-infrastructure-sectors>
8. Kaspersky Lab, State of Industrial Cybersecurity 2018 Survey. Kaspersky Lab, 2018. Available Aug 15, 2018 from Kaspersky at <https://ics.kaspersky.com/the-state-of-industrial-cybersecurity-2018/>
9. Chee-Wooi Ten, Chen-Ching Liu, and Govindarasu Manimaran, "Vulnerability assessment of cybersecurity for SCADA systems," IEEE Trans Power Systems, vol. 23, no. 4, pp. 1836-46, Nov 2008.
10. Muhammad Rizwan Asghar, György Dán, Daniele Miorandi, and Imrich Chlamtac, "Smart meter data privacy: A survey," IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2820 - 2835, 2017.
11. Witold Kinsner, "Polyscale analysis and fractional operators for cognitive systems," in Proc. 13th IEEE International Conference on Cognitive Informatics & Cognitive Computing, ICCI*CC 2014 (London, UK; August 18-20, 2014) Paper KN5, 2014.
12. Witold Kinsner, "Is entropy suitable to characterize data and signals for cognitive informatics?" Intern. J. Cognitive Informatics and Natural Intelligence, vol. 1, no. 2, pp. 34-57, Apr-Jun 2007.
13. Witold Kinsner, "A unified approach to fractal dimensions," Intern. J. Cognitive Informatics and Natural Intelligence, vol. 1, no. 4, pp. 26-46, Oct-Dec 2007.

14. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016. (ISBN: 978-026203561-3).
15. Yang Xin, Lingshuang Kong, Zhi Liu, Yuling Chen, Yanmiao Li, Hongliang Zhu, Mingcheng Gao, And Haixia Hou, And Chunhua Wang, "Machine Learning and Deep Learning Methods for Cybersecurity," *IEEE Access*, pp. 35365-81, July 19, 2018.
16. Jesus David Terrazas Gonzalez and Witold Kinsner, "Zero-crossing analysis of Lévy walks and a DDoS dataset for real-time feature extraction: Composite and applied signal analysis for strengthening the Internet-of-Things against DDoS attacks," *Int. J. Softw. Sci. Computational Intelligence IJSSCI*, vol. 8, no. 4, pp. 1–28, 2016. (doi: 10.4018/IJSSCI.2016100101)
17. Jesus David Terrazas Gonzalez and Witold Kinsner, "Zero-crossing analysis and information divergence of Lévy walks for real-time feature extraction," *Int. J. Handheld Comput. Res. IJHCR*, vol. 7, no. 4, pp. 41–59, 2016. (doi: 10.4018/IJHCR.2016100104)
18. Yosra Lakhdhar, Slim Rekhis, and Nouredine Boudriga, "Proactive security for safety and sustainability of mission critical systems," *IEEE Transactions on Sustainable Computing*, Feb 2018. doi:10.1109/TSUSC.2018.2810092)
19. Qiang Liu, Pan Li, Wentao Zhao, Wei Cai, Shui Yu and Victor C.M. Leung, "A survey on security threats and defensive techniques of machine learning: A data driven view," *IEEE Access*, vol. 6, pp. 12103-12117, 2018. (doi: 10.1109/ACCESS.2018.2805680).
20. Catriona Mairi Kennedy, *Distributed Reflective Architectures for Anomaly Detection and Autonomous Recovery*. Doctoral Thesis. Birmingham, UK: University of Birmingham, Jun 2003.
21. Awais Rashid, George Danezis, Howard Chivers, Emil Lupu, Andrew Martin, Makayla Lewis and Claudia Peersman, "Scoping the Cyber Security Body of Knowledge," *IEEE Security & Privacy*, vol. 16, no. 3, pp. 96 - 102, May/June 2018.
22. IEEE, Try-CybsI, 2018. Available Aug 15, 2018 from IEEE at <http://try.cybersecurity.ieee.org/trycybsi/>
23. ACM, ACM/IEEE/AIS SIGSEC/IFIP Cybersecurity Curricular Guideline. 2017. <http://cybered.acm.org/>
24. Hussein A. Abbass, Eleni Petraki, Kathryn Merrick, John Harvey, and Michael Barlow, "Trusted autonomy and cognitive cyber symbiosis: Open challenges," *Cognitive Computation*, vol. 8, no. 3, pp. 385–408, Jun 2018. (doi: 10.1007/s12559-015-9365-5).
25. Hussein A. Abbass, George Leu, and Kathryn Merrick, "A review of theoretical and practical challenges of trusted autonomy in big data," *IEEE Access*, vol. 4, pp. 2808-2830, June 24, 2016.

Biografia

Francesco Flammini ha ottenuto la laurea con lode in Ingegneria Informatica (2003) ed il Dottorato di Ricerca in Ingegneria Informatica e Automatica (2006) presso l'Università di Napoli Federico II. Ha avuto esperienze professionali e di ricerca sia nel mondo industriale (tra cui Ansaldo STS, Poligrafico dello Stato) che in quello accademico (tra cui University of Maryland, Linnaeus University). È attualmente membro Senior dell'IEEE, presidente del comitato tecnico IEEE SMC sulla Homeland Security, e coordinatore delle conferenze per l'iniziativa IEEE sui Symbiotic Autonomous Systems. È autore di oltre 100 pubblicazioni scientifiche su libri, riviste e atti di congressi internazionali. Il suo ultimo libro, pubblicato da Springer, si intitola "Resilience of Cyber-Physical Systems: From Risk Modeling to Threat Counteraction".

Email: flamminifra@hotmail.com



Principali sfide nello sviluppo di una soluzione SAR nel settore sanitario

Domenico Presenza, Philip Wright, Federica Saccà

Sommario

Quella dei Sistemi Autonomi Simbiotici è la visione secondo cui gli sviluppi tecnologici in corso porteranno, nel giro di poche decine di anni, all'emergere di relazioni di tipo simbiotico, comuni nel mondo animale e vegetale, anche nel contesto delle interazioni tra uomini e tecnologia. Si immagina che l'interazione tra uomini e tecnologia diventerà così stretta da portarci, probabilmente, a considerare il risultato di questa interazione come una nuova forma di vita.

È plausibile prevedere che uno dei settori dove per primo tale fenomeno avrà luogo sia quello dei sistemi per l'assistenza sanitaria. D'altronde nella letteratura l'uso del termine "simbiotico" occorre sin dai primi lavori che propongono l'uso di tecnologie robotiche per l'assistenza sanitaria.

I sistemi per l'automazione degli ambienti domestici ed i sistemi robotici possono alleviare il peso delle cure assistenziali e rendere il paradigma dell'assistenza in comunità ed in casa una realtà fornendo assistenza fisica e cognitiva agli anziani nel contesto delle attività domestiche quotidiane, aiutandoli così a mantenere la loro autonomia più a lungo e ritardando la necessità di servizi sociali e medici.

Questo articolo presenta SARA (Socially Assistive Robotic solution for Ambient assisted living), e descrive alcune delle sfide da fronteggiare nel suo sviluppo e spiega come le tecnologie IoT che offrano interoperabilità, connettività, e sicurezza end-to-end possono aiutare ad affrontare queste sfide e, quindi, contribuire alla realizzazione della visione dei Sistemi Autonomi Simbiotici.

Keywords: Socially Assistive Robotics, Ambient Assisted Living, Healthcare, IoT, Artificial Intelligence

Introduzione

Quella dei Sistemi Autonomi Simbiotici (Symbiotic Autonomous Systems) [1] è la visione secondo cui gli sviluppi tecnologici in corso porteranno, nel giro di poche decine di anni, all'emergere di relazioni di tipo simbiotico, comuni nel mondo animale e vegetale, anche nel contesto delle interazioni tra uomini e tecnologia. Si immagina che l'interazione tra uomini e tecnologia diventerà così stretta da portarci, probabilmente, a considerare il risultato di questa interazione come una nuova forma di vita.

Se questa visione si realizzerà lo farà in una società certamente molto più "anziana" di quella odierna [2]. È plausibile prevedere, dunque, che uno dei settori dove per primo tale fenomeno avrà luogo sia quello dei sistemi per l'assistenza sanitaria [1]. D'altronde nella letteratura l'uso del termine "simbiotico" occorre sin dai primi lavori che propongono l'uso di tecnologie robotiche per all'assistenza sanitaria [3].

Questo articolo presenta alcuni degli ostacoli tecnologici che ancora occorre superare nell'immediato prima di poter realizzare completamente la visione SAS nel dominio dell'assistenza sanitaria.

La prospettiva legata alla tendenza globale di un aumento delle aspettative di vita è che tale tendenza sarà accompagnata da un rapido incremento del numero di persone affette da forme moderate di demenza [4] o morbo di Alzheimer [5] [6]. L'invecchiamento e la demenza portano ad una perdita delle capacità fisico-cognitive cronica e significativa e ad un declino delle capacità funzionali che si traducono in una aumentata dipendenza e necessità di cure assistenziali [7], con rilevanti oneri medici, sociali, psicologici e finanziari che gravano sui pazienti, le loro famiglie e le rispettive comunità di appartenenza [8]. In effetti, il tipo di cura continuo, onnicomprensivo e integrato necessario alle persone affette da demenza non può essere soddisfatto dall'odierno modello specialistico [9] e, inoltre, la differenza tra il numero di persone bisognose di cura e quelle in grado di offrirle è destinata ad aumentare [10]. Inoltre, quando i problemi legati all'invecchiamento cominciano ad apparire, le persone preferiscono continuare a vivere nella propria casa il più a lungo possibile piuttosto che essere istituzionalizzate in residenze assistenziali [11]. Questo ha portato a proporre nuovi paradigmi di assistenza più accettabili e sostenibili come, ad esempio, quello dello "Aging in place" [12] in cui una persona anziana può scegliere di vivere in un ambiente (luogo/comunità) meno vincolante e, soprattutto, di sua scelta.

I sistemi per l'automazione degli ambienti domestici ed i sistemi robotici possono alleviare il peso delle cure assistenziali e rendere il paradigma dell'assistenza in comunità ed in casa una realtà fornendo assistenza fisica e cognitiva agli anziani nel contesto delle attività domestiche quotidiane, aiutandoli così a mantenere la loro autonomia più a lungo e ritardando la necessità servizi sociali e medici [13]. I progressi nelle tecnologie dell'intelligenza artificiale e dell'Internet delle cose (IoT) stanno convergendo rapidamente rendendo tali soluzioni fattibili sia tecnicamente che economicamente. In questa prospettiva, Engineering [14] sta sperimentando

l'estensione della sua offerta commerciale AREAS® - un sistema ERP per il settore della Sanità - con SARA (Socially Assistive Robotic solution for Ambient assisted living), un sistema robotico che mira a fornire un sistema automatizzato di monitoraggio e assistenza domestica alle persone anziane con forme moderate di demenza. Questo articolo presenta SARA, descrive alcune delle sfide da fronteggiare nel suo sviluppo e spiega come le tecnologie IoT che offrano interoperabilità, connettività, e sicurezza end-to-end possono aiutare ad affrontare queste sfide e, quindi, contribuire alla realizzazione della visione dei Sistemi Autonomi Simbiotici.

Questo articolo è organizzato come segue: la sezione 1 presenta una panoramica del sistema; la Sezione 2 presenta uno scenario applicativo e riporta alcuni dei compiti assistivi coadiuvati da SARA; la Sezione 3 descrive le sfide tecnologiche nello sviluppo di SARA e la sezione 4 fornisce le conclusioni.

1. Il sistema SARA

Come annunciato dal nome, SARA adotta il paradigma della Socially Assistive Robotics (SAR) che [16] si colloca all'intersezione tra la Assistive Robotics (AR) - robot che forniscono assistenza fisica (es. robot per la riabilitazione, supporti alla mobilità, bracci meccanici) - e Socially Interactive Robotics (SIR) - robot il cui compito principale è di stabilire una qualche forma di interazione sociale con gli esseri umani (es. ingaggiando una conversazione, comprendendo i gesti, assumendo un ruolo, collaborando alla realizzazione di un compito). Il paradigma SAR mira a fornire assistenza agli esseri umani attraverso interazioni sociali. Un obiettivo ambizioso considerato che probabilmente il più complesso di tutti i comportamenti umani. Un robot che abbia competenze di tipo sociale dovrebbe per esempio [17]: comunicare attraverso un dialogo di alto livello; utilizzare segnali non verbali (sguardo, gesti, ecc.), esprimere e percepire emozioni; apprendere/riconoscere modelli di altri agenti; mostrare un carattere ed una personalità proprie; stabilire e mantenere relazioni sociali; apprendere/sviluppare competenze sociali.

Raggiungere in tali aree un livello di competenza paragonabile a quello umano è ben oltre anche le migliori tecnologie AI odierne e, dunque, gli obiettivi di SARA sono molto più pragmatici. L'intento è, dunque, quello di sviluppare una piattaforma espandibile che inizialmente mostri capacità sociali basilari (ben dentro i limiti delle attuali tecnologie), ma che, allo stesso tempo, rappresenti una solida base per future espansioni verso forme di interazione più sofisticate e più simili a quelle umane.

In linea con precedenti lavori sugli "ambienti intelligenti" adatti agli anziani [13] [18], il progetto di SARA integra un numero di tecnologie assistive per affrontare i diversi problemi che gli anziani fronteggiano quando vivono indipendentemente e con il supporto della loro comunità di appartenenza. Più specificatamente, il sistema include vari elementi robotici (sia statici che mobili) che lavorano in collaborazione con dispositivi indossabili, smart devices personali (telefoni o tablets) e vari servizi computazionali (proprietary e di terze parti) che continuamente monitorano e assistono i soggetti anziani (e i loro

badanti) nelle loro normali attività quotidiane all'interno e intorno la loro casa. I componenti tecnologici di SARA sono:

- servizi cloud AREAS® (ACS): servono principalmente come repository di informazioni cliniche e per la persistenza dei dati biomedici raccolti dalle funzioni di monitoraggio.
- una Body Area Network (BAN) comprendente periferiche indossabili per la raccolta dei dati biomedici (es. battito cardiaco, pressione sanguigna, frequenza respiratoria, livelli di stress, bilanciamento e cadute) indossati dai soggetti anziani e che comunicano via Bluetooth con altri dispositivi personali (telefoni o tablets)
- un Deambulatore Robotico (RR): un carrellino motorizzato che fornisce supporto durante la deambulazione ed in grado di monitorare la postura e l'andatura del suo utilizzatore e, con una forma limitata di capacità di navigazione autonoma.
- un Assistente Robotico (RA): un robot umanoide - specificatamente un robot Pepper di Softbank [19] - capace di navigazione autonoma, riconoscimento dei volti, dialoghi predefiniti e assumere atteggiamenti simili a quelli umani (es. gesticolare)
- vari servizi di AI. risorse computazionali che forniscono, ad esempio, meccanismi di apprendimento automatico, mappatura, pianificazione di percorsi e ragionamento.
- una Smart House: una abitazione che contiene oggetti "intelligenti" che utilizzano reti cablate o senza fili e che hanno la capacità di analizzare pattern di attività, di gestire anticipando le necessità degli abitanti



Figura 1
Componenti e comunicazioni di SARA

Con l'eccezione dei componenti ACS e RR (entrambi sviluppati da Engineering) tali componenti sono tutti componenti off-the-shelf che non sono stati progettati specificatamente per lavorare insieme. I vari componenti producono e consumano differenti quantità e tipi di dati, in diversi formati, supportano diversi protocolli di comunicazione (Bluetooth, Wi-Fi, Ethernet, can-bus, ecc.). Inoltre

per ragioni di affidabilità i dati critici sono collezionati da più sorgenti in modo ridondante (più componenti possono monitorare le stesse caratteristiche) e condivisi attraverso più canali anch'essi ridondanti. Per assicurare una tempestiva risposta dei robots, gli algoritmi di AI che richiedono significative risorse di calcolo sono eseguite concorrentemente nel sistema, con operazioni che vengono distribuite nei nodi di calcolo più prossimi. In SARA le configurazioni di rete non sono statiche: i dispositivi mobili possono uscire dalla copertura Wi-Fi; i nodi collegati alla rete di alimentazione (es. stazioni base del Wi-Fi, server domotici) possono essere disconnesse a causa di interruzioni nella fornitura di elettricità, mentre quelli alimentati da batterie necessitano di ricariche periodiche. Le periferiche indossabili (BAN) possono essere rimosse per varie ragioni (es. per una doccia).

In breve, già solo in termini di hardware, software, struttura e comunicazioni di rete SARA si configura come un sistema eterogeneo, distribuito e dinamico che pone un numero significativo di sfide. Ulteriori sfide, come descritto nella sezione successiva, derivano dalle peculiarità dello scenario applicativo.

2. Lo scenario applicativo: assistenza domiciliare remota agli anziani

Come già evidenziato SARA è una soluzione robotica volta a fornire sia un servizio di monitoraggio delle condizioni di salute e di eventuali incidenti sia a fornire assistenza agli anziani con forme moderate di demenza. La soluzione è intesa per essere fornita come servizio, con una appropriata supervisione, dalle aziende sanitarie ai cittadini eleggibili per la fornitura del servizio.

I servizi di assistenza forniti da SARA includono: tener traccia della posizione degli oggetti importanti per il paziente (es. una confezione di medicinali); guidare e fornire supporto alla mobilità agli anziani quando si muovono nel loro ambiente; monitorare le condizioni di salute (es. il BAN rileva una perdita di equilibrio, il deambulatore rileva un cambiamento nell'andatura, il Robotic Assistant si accorge di un cambiamento dell'espressione); notificare a chi si prende cura del paziente un possibile problema o incidente; consentire al medico o al badante di vedere come sta un paziente; supportare esercizi di riabilitazione fisica o cognitiva; ricordare al paziente di prendere le medicine quando deve; memorizzare e fornire accesso ai dati medici monitorati dal sistema.

Lo scenario applicativo immaginato per SARA rende evidente l'intimo e critico ruolo che il sistema gioca nella vita dell'anziano. Il sistema è integrato nella abitazione e, per via dell'utilizzo continuo di dati personali e sensibili che riguardano le attività in corso ed lo stato di salute entra a far parte in modo intimo della vita quotidiana della persona assistita. Questo naturalmente pone requisiti stringenti di sicurezza e privacy che il sistema SARA deve soddisfare. Il ruolo del sistema è anche critico rispetto agli aspetti della salute: fa la differenza se il sistema fallisce nell'individuare un incidente relativo alla salute, o di ricordare al paziente di prendere le medicine. Pertanto SARA deve essere robusto ed affidabile, il che pone ulteriori sfide tecnologiche, specialmente in considerazione

della sua natura di sistema eterogeneo, dinamico e distribuito. Infine, come per ogni automatismo presente in una casa, e tanto più se di un anziano, SARA non deve rappresentare un rischio per l'incolumità dei suoi abitanti.

In definitiva appare chiaro come la realizzazione di SARA pone sfide tecnologiche fondamentali, sin da subito, anche quando ci si pone l'obiettivo di abilitare interazioni sociali elementari.

3. Le sfide tecnologiche

Questa sezione descrive le sfide tecnologiche poste dalla progettazione e dalla realizzazione del sistema SARA con particolare attenzione a quegli aspetti dove le innovazioni riguardanti l'interoperabilità, la connettività e sicurezza dei sistemi IoT possono svolgere un ruolo chiave nella realizzazione della soluzione.

3.1 Sicurezza, privacy, affidabilità e incolumità

Nella precedente sezione abbiamo evidenziato la forte necessità di soddisfare requisiti di sicurezza, privacy, affidabilità e incolumità. Questa importanza non sarà mai ribadita abbastanza. SARA è una tecnologia indirizzata ad uno specifico settore del dominio sanitario. Il suo scopo è fornire reali benefici ai cittadini anziani e ai loro badanti alleviando il carico da servizi sociali e sanitari sovraccarichi. Questo obiettivo complessivo non dovrebbe essere inficiato dalla tecnologia stessa. In particolare, il sistema deve essere in sicurezza rispetto ad intrusioni, deve garantire la privacy dei cittadini anziani e dei loro badanti, in particolare modo per quel che riguarda la trasmissione e la memorizzazione di dati che riguardano le condizioni di salute, e deve, per offrire reali benefici, eseguire le sue funzioni in modo affidabile e sicuro.

I requisiti di sicurezza, privacy ed incolumità includono: tutti i dati personali memorizzati (es. dati medici, dati raccolti dal BAN) devono essere criptati ed accessibili solo a persone o processi opportunamente autenticati ed autorizzati; tutte le trasmissioni contenenti dati personali, incluse le notifiche e i flussi audio/video trasmessi nel contesto delle sessioni di tele-presenza, devono essere criptate; il trattamento dei dati medici dei pazienti deve essere conforme agli standard del settore (es. [21]); tutti i segnali di controllo (sia che viaggino all'interno di un device o sulla rete) devono essere protetti da interferenze intenzionali (es. non si vuole che un hacker possa prendere il controllo dello RR o del RA; i componenti e le tecnologie del sistema devono essere configurate su una rete affidabile per evitare interruzioni nel servizio; le trasmissioni con alta priorità e.s. la segnalazione di possibile evento di caduta, devono essere trasmesse in modo affidabile (es. trasmettendo lo stesso messaggio su più canali) e con una opportuna priorità di rete; il sistema SARA, ed in particolare le sue componenti robotiche, devono includere misure di protezione che garantiscano di non compromettere l'incolumità dei pazienti (o chiunque altro); i componenti del sistema e le tecnologie di comunicazione impiegati non devono interferire o danneggiare le apparecchiature medicali (specularmente, per ragioni di affidabilità, non dovrebbero essere influenzati dalla presenza di dispositivi medicali).

Garantire il contemporaneo soddisfacimento di tutti i requisiti sopra elencati è di per sé una sfida. Quello di cui occorre disporre sono strumenti per valutare il modo accurato come cambiano le proprietà di sicurezza, privacy, affidabilità e incolumità al variare della configurazione del sistema ad esempio per valutare se usando una periferica, o un gruppo di periferiche, per un particolare scopo può esporre il sistema ad un livello di rischio inaccettabile.

Per quanto detto sopra esiste dunque la necessità di una infrastruttura IoT che offra la possibilità di verificare in modo formale le proprietà complessive di sicurezza, privacy, affidabilità e incolumità di una applicazione basata su di essa.

3.2 Interoperabilità

Come descritto nella sezione 2, il sistema SARA si basa su una rete di sensori e attuatori che impiegano varie tecnologie di comunicazione sia wireless che cablate. I nodi nella rete sono eterogenei e includono da singoli sensori quali IMU o sensori di presenza a sofisticati componenti robotici con la propria rete di sensori/attuatori e computer di bordo. Questa rete di sensori e attuatori comunica anche con servizi cloud di back-end per memorizzare e recuperare eventi (es. "evento caduta") e dati provenienti dalle attività di monitoraggio (es. serie temporali di misure rilevate dal BAN).

Conseguentemente il sistema deve essere in grado di gestire un ampio repertorio di periferiche, formati dati (rappresentazioni sintattiche), conversioni di unità di misura (per le letture dei sensori) e protocolli di comunicazione (es. WiFi, ZigBee, Bluetooth). Inoltre vari aspetti chiave delle funzionalità di SARA richiedono l'aggregazione e l'analisi di dati raccolti da molteplici fonti. L'individuazione affidabile di un evento di caduta, per esempio, può richiedere il confronto continuo dei dati raccolti dalle IMU del BAN, i sensori di pressione montati nelle maniglie del RR, e le videocamere a bordo del RA. In breve il sistema SARA deve affrontare il problema della interoperabilità dei dati e delle periferiche IoT.

I problemi di interoperabilità semantica si manifestano in tutte le fasi del ciclo di vita del sistema e, pertanto, è necessario disporre di meccanismi di interoperabilità sintattica e semantica:

- durante la progettazione per gestire l'eterogeneo insieme di protocolli di comunicazione ed applicativi necessariamente introdotti nel sistema dall'utilizzo di sensori ed attuatori di serie disponibili sul mercato, utilizzo di piattaforme robotiche (es ROS [22]) e di sistemi demotici proprietari;
- nella fase di configurazione per far corrispondere alle specifiche necessità indicate dai medici (es. il paziente necessita di monitoraggio cardiaco) configurazioni personalizzate per ciascun paziente che richiedono l'integrazione di periferiche non previste durante la fase di progettazione del sistema
- nella fase operativa per abilitare l'integrazione in modo opportunistico di servizi di terze parti (es servizi di AI) e componenti IoT (es. un paramedico che interviene per gestire un incidente può avere la necessità di connettere al sistema una attrezzatura medica particolare e trasmettere i dati raccolti ad un medico che fornisce assistenza remota)

Come esempio specifico, i meccanismi di interoperabilità potrebbero essere usati per modellare, attraverso un formato standard: i) la semantica e le capacità dei devices a bordo del deambulatore robotico (RR) ii) descrivere i requisiti di una applicazione di monitoraggio necessaria ad uno specifico paziente (es. analisi dell'andatura) il che consentirebbe di dispiegare una nuova funzionalità di analisi dell'andatura su un deambulatore inizialmente non progettato per tale scopo. Questo grazie alla possibilità di far corrispondere, in maniera automatica, la descrizione semantica delle periferiche a bordo del carrellino con i requisiti specifici del paziente.

In conclusione, per poter ottenere una corretta integrazione funzionale delle vasta varietà di componenti e interazioni all'interno di SARA si richiede un elevato grado di interoperabilità sintattica e semantica.

3.3 Intelligenza Artificiale

SARA applicando il paradigma della Socially Assistive Robotics, si caratterizza fondamentalmente come un sistema di intelligenza artificiale. Raggiungere capacità di interazione sociale che siano al livello di quelle umane, come già evidenziato nella introduzione, è oltre le capacità odierne della AI. Tuttavia ha senso ambire ad ottenere qualche forma primitiva di competenza sociale e riuscire a costruire una base solida per la definizione di competenze di più alto livello. In particolare i recenti sviluppi nel campo dell'apprendimento automatico (ML) e la crescente disponibilità di ampi insiemi di dati per l'addestramento di sistemi di ML ci consente di affrontare alcuni dei problemi fondamentali della classificazione di oggetti in base alla percezione sensoriale che se ne ha.

Soluzioni ML per il riconoscimento di oggetti, persone/volti e parlato, per esempio, sono già disponibili (e sono incorporate nelle più moderne periferiche mobili). Tecnologie per il riconoscimento delle emozioni umane (es. Eyeris [23]) e i gesti delle mani (es. uSens FINGO [24]) stanno aparendo sul mercato e, anche i prodotti più vecchi come la Kinect di Microsoft [25] sono in grado di fornire il riconoscimento di gesti e postura, sebbene in modo grossolano [26]. La disponibilità di enormi quantità di testi sul web ha anche prodotto progressi nella elaborazione del linguaggio naturale (NLP) con tecniche di ML come evidenziato da Google Assistant [28], Siri di Apple [29] e IBM Watson [30] che combinano le tecniche di NLP con conoscenza/esperienza di dominio per fornire risposte sensate a domande poste in linguaggio naturale. Tali sistemi stanno gradualmente evolvendo in agenti software in grado di conversare [31], capaci (attualmente in forma limitata) di sostenere un discorso su un tema specifico. Esistono anche agenti software con diversi caratteri ("personas") [32] per sostenere una conversazione ancora più naturale.

Per creare interazioni sociali sofisticate avremo la necessità di integrare tutti questi approcci per essere in grado di riconoscere simultaneamente oggetti e persone, parlato e linguaggio gestuale, postura e andatura, per arricchire queste classificazioni degli aspetti percettivi con conoscenza di dominio e di orchestrare il tutto con la gestione della conversazione e del dialogo. Tutto questo comporta un carico computazionale da sostenere sottostando a vincoli temporali stringenti.

L'apprendimento automatico è anche una tecnologia chiave per la navigazione autonoma. Le auto senza guidatore di Google [33] e Apple [34], per esempio, dipendono in maniera critica da algoritmi di deep learning che operano su nuvole di punti tridimensionali ottenuti tramite sensori LIDAR (Light Detection and Ranging) per distinguere passaggi e ostacoli e identificare e tracciare oggetti (es. altri veicoli) e persone. Recentemente, sia Google che Apple hanno rilasciato API per la realtà aumentata su dispositivi mobili che includono funzioni di ML per l'identificazione e il tracciamento di superfici verticali/orizzontali e oggetti all'interno di flussi video. Queste API ed altre tecniche adatte ad essere disponibili anche su telefoni mobili (es. [35]) possono essere utilizzate per affrontare il problema della localizzazione e mappatura in simultanea (SLAM [36]): cioè il problema di mappare e localizzare la posizione di un agente in un ambiente sconosciuto. Spostandosi dagli aspetti della percezione ai sistemi per la generazione di azioni, le reti neurali sono usate da lungo tempo per controllare i motori di robot con ruote (es. [37] per un esempio piuttosto recente), e più recentemente per la simulazione dei movimenti di corsi rigidi più complessi, quali ad esempio i movimenti coordinati richiesti per la camminata, il galoppo e la corsa dei robot quadrupedi di DARPA/Boston Dynamics [38]. Reti non supervisionate basate sul principio di minimizzazione dell'errore di predizione (PEM) [39] (es. discendenti delle macchine di Helmholtz [38]) sembrano promettenti (in un futuro non troppo lontano) per lo sviluppo di complessi sistemi di percezione-azione. Di nuovo, questi compiti percettivi e attuativi sono computazionalmente onerosi e devono soddisfare vincoli temporali stringenti.

Come ribadito all'inizio, le interazioni sociali sono senza dubbio i più complessi dei comportamenti umani, e la comunità che si occupa di robot autonomi è ancora molto lontana dal riprodurre il livello di competenza sociale paragonabile a quello umano. Le tecniche di ML, come brevemente ricordato sopra, hanno fatto notevoli passi verso questo obiettivo, primariamente nell'area della percezione producendo quel tipo di input sensoriale sul quale, in futuro, potrebbero essere costruiti comportamenti sociali sofisticati. Di fatto i progressi nelle tecnologie delle reti neurali (quali PEM) sembrano pronti anche per affrontare i più ampi aspetti comportamentali della AI. Il punto chiave, comunque, è che SARA richiederà come minimo il continuo operare di un repertorio di sistemi percettivi. Questo implica la necessità di molteplici processi di AI/ML che lavorano in parallelo per estrarre in modo tempestivo, senza latenze di rete significative, differenti caratteristiche dagli input sensoriali. Ciò ci consente di identificare la forte necessità di una infrastruttura IoT che supporti elaborazioni di AI computazionalmente onerose.

3.4 Gestione della rete

SARA è un'applicazione distribuita con differenti entità computazionali, in grado di comunicare in modo continuativo e di sincronizzare le proprie attività per generare le funzionalità complessive del sistema. I fattori che sottolineano la necessità di una distribuzione computazionale di SARA sono: la necessità di risorse computazionali ai nodi, per eseguire calcoli complessi quali, ad esempio, riconoscimento di oggetti/persone e navigazione autonoma; l'esigenza

di ridurre l'elaborazione dei calcoli per il risparmio energetico dei dispositivi (ad esempio il RR e il RA); la necessità di condividere le conoscenze tra i dispositivi (ad esempio per gli eventi rilevati dai dispositivi domotici che devono essere gestiti da quelli robotici) o di conoscenze condivise sul cloud (ad esempio, fonti di dati remoti); la richiesta della ridondanza su varie parti del sistema, per garantire l'affidabilità e la sicurezza.

Allo stesso tempo, molte di queste operazioni computazionali, a causa della richiesta prossima al tempo reale, sono soggette a vincoli infrastrutturali tali da garantire una bassa latenza e un'alta affidabilità di comunicazione tra i vari componenti.

Un altro aspetto rilevante, è che tutti i nodi robotici di SARA, hanno comportamenti (quasi)autonomi guidati da eventi incerti e imprevedibili nell'ambiente in cui si trovano, con una corrispondente incertezza ed imprevedibilità nel carico di calcolo generato.

Tale variabilità, impedisce la determinazione anticipata di una configurazione ottimale delle risorse di rete. Per tale motivo, è necessaria una variabilità del meccanismo che comporta l'adozione di sistemi opportunamente dinamici e flessibili per la gestione della rete. Questa è precisamente il tipo di flessibilità offerta dalla tecnologia delle Software Defined Network (SDN) [41], che rende possibile configurare dinamicamente, tramite software, sia il corretto percorso dei dati nella rete che le interfacce di comunicazione alternative in funzione dello stato corrente della rete e delle esigenze QoS (Quality-of-Service, qualità del servizio) dell'applicazione. Abbiamo dunque identificato una forte esigenza in SARA di soluzioni flessibili e adattive di gestione della rete, quale SDN.

4. Conclusioni

In questo documento abbiamo descritto SARA, una soluzione di tipo SAR (Socially Assistive Robotics), attualmente in fase di sviluppo, finalizzata a fornire il monitoraggio automatico degli incidenti e dei dati clinici per l'assistenza domiciliare agli anziani con deficit cognitivi lievi. Abbiamo delineato i componenti tecnici di SARA e illustrato la sua applicazione prevista: in particolare, la sua natura altamente eterogenea, dinamica, distribuita, funzionale e comportamentale nel suo complesso. Sebbene i recenti progressi nella IA e nella IoT (insieme ai costi decrescenti dell'hardware e all'ampia disponibilità di risorse di cloud computing) stiano convergendo rendendo le soluzioni SAR sempre più tecnicamente ed economicamente fattibili, il loro sviluppo deve ancora affrontare notevoli sfide tecnologiche. In particolare, abbiamo identificato quattro sfide principali in SARA, per le quali l'interoperabilità, la connettività e la sicurezza dei sistemi IoT, potrebbero risultare di notevole valore. Queste sfide sono: i) forti requisiti di sicurezza, privacy, affidabilità ii) interoperabilità di dispositivi, e di dati, affrontata con meccanismi di trasformazione semantica, iii) sfide nell'intelligenza artificiale, evidenziando la necessità di spostare i processi di ML sui dispositivi IoT e iv) sfide nella gestione della rete, dove abbiamo identificato la necessità di una soluzione flessibile ed adattiva quale quella offerta dalle SDN.

Ringraziamenti

Gli autori vogliono ringraziare il supporto offerto dal progetto “SemIoTics”, identificativo numero 780315, del programma Horizon 2020. Titolo per esteso del progetto: Smart End-to-end Massive IoT Interoperability, Connectivity and Security.

Riferimenti

1. S. Mason Dambrot, D. de Kerchove, F. Flammini, W. Kinsner, L. MacDonald Glenn, R. Saracco, “Symbiotic Autonomous Systems White Paper II”, October 2018 <https://www.diva-portal.org/smash/get/diva2:1260812/FULLTEXT02.pdf>
2. R. Suzman and J. Beard, “Global Health and Aging”, NIH Publ. no 117737, Vol. 1, no. 4, pp. 273-277, 2011
3. T. Sato, Y. Nishida, H. Mizoguchi, “Robotic room: Symbiosis with human through behavior media” Robotics and Autonomous Systems Volume 18, Issues 1–2, July 1996, Pages 185-194
4. Mild Cognitive Impairment (MIC) - <https://www.alz.org/dementia/mild-cognitive-impairment-mci.asp>
5. Alzheimer disease - https://www.alz.org/alzheimers_disease_what_is_alzheimers.asp
6. Alzheimer disease in World Health Organization - <http://www.who.int/en/news-room/fact-sheets/detail/mental-health-of-older-adults>
7. R. Brookmeyer, E. Johnson, K. Ziegler-graham, and H. M. Arrighi, “Forecasting the Global Burden of Alzheimer’s Disease”, Sel. Work., vol. 130
8. D.P. Rice, P.J. Fox, W. Max, et al., “The economic burden of Alzheimer’s disease care”, Health Aff (Millwood), 1993;12:164-176
9. M. Prince, A. Comas-Herrera, M. Knapp, M. Guerchet, and M. Karagiannidou, “World Alzheimer Report 2016 Improving healthcare for people living with dementia. Coverage, Quality and costs now and in the future,” Alzheimer’s Dis. Int., pp. 1–140, 2016.
10. N. Super, and J. D. Moore, “Who will be there to care? The growing gap between caregiver supply and demand”, 2002.
11. J. Broekens, M. Heerink, and H. Rosendal, “Assistive social robots in elderly care: a review”, Gerontechnology, vol. 8, no. 2, 2009.
12. K.D. Marek, and M.J. Rantz “Ageing in place: a new model for long-term care”, Nurs Admin Quart 24:1–116, 2000.
13. C. Röcker, “Intelligent environments as a promising solution for addressing current demographic changes”, Int. J. Innov. Technol. Manag., vol. 4, no. 1, pp. 76–79, 2013. Available at: <http://www.ijimt.org/papers/361-D012.pdf>
14. Engineering Ingegneria Informatica spa - <https://www.eng.it>
15. SEMIoTIC project - <https://www.semiotics-project.eu>

16. D. Feil-Seifer, and M. J. Matarić, "Defining socially assistive robotics", in proceedings of the 2005 IEEE 9th International Conference on Rehabilitation Robotics June 28 - July 1, 2005, Chicago, IL, USA, pp. 465-468
17. T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots", in Robotics and Autonomous Systems, vol.42(3-4) pp. 143- 166, 2003
18. E. Kartsakli, et al., "A survey on M2M Systems for mHealth: A wireless communications perspective", MDPI Sensors, 2014
19. Pepper robot - <https://www.softbankrobotics.com/emea/en/robots/pepper>
20. HL7 standards - <http://www.hl7.org/>
21. General Data Protection Regulation - <https://gdpr-info.eu/>
22. ROS - www.ros.org
23. Eyeris - <https://www.cnet.com/roadshow/news/eyeris-emovu-detects-driver-emotions/>
24. Fingo - <https://en.usens.com/products/fingo/>
25. Kinect - <https://en.wikipedia.org/wiki/Kinect>
26. M. Gabel, R. Gilad-Bachrach, E. Renshaw and A. Schuster, "Full body gait analysis with kinect", in Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE Engineering in Medicine and Biology Society, vol. 8, pp. 1964-7, 2012 [<https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/KinectGaitEMBC2012.pdf>, April 2018]
27. Y. Goldberg, "A primer on neural network models for natural language processing", in Journal of Artificial Intelligence Research, vol. 57, pp. 345-420, 2016 [<https://www.jair.org/media/4992/live-4992-9623-jair.pdf> , April 2018]
28. Google Assistant - <https://assistant.google.com/>
29. Apple's Siri - <https://www.apple.com/ios/siri/>
30. IBM's Watson - <https://www.ibm.com/watson/>
31. V. Oriol, and V. de Quoc, "A neural conversational model", in Proceedings of the 31st International Conference on Machine Learning, Lille, France, 2015, vol. 37 [<https://arxiv.org/pdf/1506.05869v1.pdf>]
32. J. Li, M. Galley, C. Brockett, G.P. Spithourakis, J. Gao, and B. Dollan, "A persona-based neural conversation model", [the arXiv preprint repository], 2016 [<https://arxiv.org/pdf/1603.06155.pdf>]
33. Waymo - <https://waymo.com/tech/>
34. Y. Zhou, and O. Tuzel: "VoxelNet: End-to-end learning for point cloud based 3D object detection", [the arXiv preprint repository], 2016 [<https://arxiv.org/pdf/1711.06396.pdf>]

35. R.M. Faragher, and R.K. Harle, "SmartSLAM – An efficient smartphone indoor positioning system exploiting machine learning and opportunistic sensing", vol. 2, pp. 1006-1019, 2013 [<https://www.cl.cam.ac.uk/~rmf25/papers/SmartSLAM1.pdf>]
36. SLAM - https://en.wikipedia.org/wiki/Simultaneous_localization_and_mapping
37. J.A. Fernandez-Leon, G.G. Acosta, and M.A. Mayosky, "Behavioural control through evolutionary neurocontrollers for autonomous mobile robot navigation", in Robotics and Autonomous Systems, vol. 57, pp. 411-419, 2009
38. J.W. Lee, "A machine-learning approach to developing rigid-body dynamics simulators for quadruped trot gaits" [<http://cs229.stanford.edu/proj2008/Lee-AMachineLearningApproachToDevelopingRigid-bodyDynamicsSimulatorsForQuadrupedTrotGaits.pdf>]
39. Andy Clark, "Surfing Uncertainty: Prediction, Action and the Embodied Mind", Oxford University Press, 2016
40. G. E. Hinton, P. Dayan, B. J. Frey, and R. Neal, "The wake-sleep algorithm for unsupervised Neural Networks". Science, 268, pp 1158-1161. 1995
41. T.D. Nadeau, and K. Gray, "SDN: Software Defined Networks: an authoritative review of network programmability technologies", O'Reilly Media, Inc., 2013

Biografie

Domenico Presenza ha conseguito la laurea in Scienze dell'Informazione presso l'Università di Pisa nel 1989. Dal 2000 lavora presso la direzione Ricerca & Innovazione di Engineering Ingegneria Informatica S.p.A. presso cui attualmente guida un gruppo di lavoro sui sistemi socio-tecnici e lo sviluppo della soluzione SARA. I suoi interessi attuali riguardano principalmente il calcolo distribuito, i sistemi multi-agente, e la socially assistive robotics.

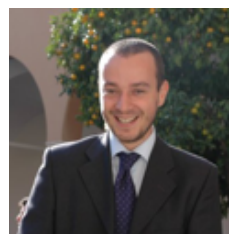
Email: domenico.presenza@eng.it



Philip Wright ha conseguito una laurea ingegneria meccanica presso l'Università degli Studi di Roma "La Sapienza". Dall'Aprile del 2013 fa parte del gruppo di ricerca e sviluppo di Engineering Ingegneria Informatica Spa come ricercatore. Ha partecipato a diversi progetti di ricerca Europei e nazionali.

Attualmente si occupa dell'integrazione dei componenti robotici e della progettazione degli aspetti meccanici della soluzione SARA.

Email: philip.wright@eng.it



Federica Saccà ha conseguito la laurea triennale in Ingegneria Informatica presso l'Università degli Studi della Calabria e la laurea specialistica in Ingegneria Biomedica presso l'Università Magna Graecia di Catanzaro. Ha successivamente approfondito le sue conoscenze in campo biomedico frequentando il Master di II livello in Ingegneria Clinica presso l'Università degli Studi di Firenze. Dal 2016 lavora come ricercatrice nel gruppo di Ricerca e Sviluppo di Engineering Ingegneria Informatica S.p.A. Ha partecipato a diversi progetti di ricerca Europei e Nazionali ed è co-autrice di diverse pubblicazioni scientifiche. Attualmente si occupa dello sviluppo di soluzioni per il miglioramento dell'assistenza sanitaria.

Email: federica.sacca@eng.it



Fra Pubblico e Privato: Sistemi Digitali Simbiotici per la Persona

Paolo Traverso, Marco Pistore, Elisabetta Farella, Amy L. Murphy, Annapaola Marconi, Pamela Forner

Sommario

In questo articolo propongo una visione alternativa di Sistema Autonomo Simbiotico, visione che è ispirata ad un importante obiettivo alla base dei sistemi simbiotici, ovvero quello di costruire soluzioni digitali in grado di trasformare la vita delle persone, una tecnologia che migliori la qualità della loro vita. Questa visione pone un obiettivo che può essere raggiunto solo con un coinvolgimento profondo e costante delle persone per le quali si vuole costruire il sistema simbiotico, un coinvolgimento ex-ante, ovvero nel processo di definizione del sistema digitale, secondo un paradigma di forte "connessione" della tecnologia digitale con le persone.

Questa visione assume concretezza nelle attività svolte da FBK insieme alla Provincia Autonoma di Trento di cui alcune presentate in questo articolo.

In questo breve articolo, vorrei proporre una visione alternativa di Sistema Autonomo Simbiotico, una visione alternativa che è ispirata ad un importante obiettivo alla base dei sistemi simbiotici, ovvero quello di costruire soluzioni digitali in grado di trasformare la vita delle persone, una tecnologia che migliori la qualità della loro vita. A mio parere, questa visione pone un obiettivo che può essere raggiunto solo con un coinvolgimento profondo e costante delle persone per le quali si vuole costruire il sistema simbiotico, un coinvolgimento

ex-ante, ovvero nel processo di definizione del sistema digitale, secondo un paradigma di forte "connessione" della tecnologia digitale con le persone.

Questa "connessione" può essere attuata solo se non ci si ferma agli aspetti tecnologici, ma il processo di costruzione del sistema simbiotico riesce a coprire tutta la "filiera" che parte sì dalla tecnologia, ad esempio da sensori, attuatori e componenti sempre più autonomi e intelligenti, dall'internet delle cose, dalla raccolta massiva di dati dall'ambiente e dalle persone stesse, dalla loro analisi

tramite le moderne tecniche di apprendimento automatico (il cosiddetto *"machine-deep-learning"*), dalla generazione di servizi e applicazioni a partire da questi dati e dalla loro analisi ... ma non si ferma lì, per arrivare a coinvolgere le persone come parte del processo di costruzione del sistema simbiotico.

Anzi, sono profondamente convinto che ridurre l'Intelligenza Artificiale al mero apprendimento automatico (profondo o meno che sia), ai meri modelli predittivi e analisi di dati sia riduttivo. La vera intelligenza artificiale è proprio la scienza che copre tutta la filiera, fino alla persona, fino ad essere una intelligenza artificiale *"human aware"*.

La visione dei *"sistemi digitali simbiotici per la persona"*, oltre a coprire tutta la filiera dalla tecnologia alla persona, deve considerare diversi attori e portare loro un valore aggiunto: le persone che hanno partecipato attivamente alla definizione del sistema simbiotico, la pubblica amministrazione, i *"policy makers"*, che possono promuovere e usufruire di *"digital twins"* del territorio, *"digital companions"* che li aiutano a definire le strategie per rendere migliore la vita delle persone che vivono sul territorio; le aziende possono trovare nelle esigenze dei cittadini e della pubblica amministrazione un mercato per il loro business; i ricercatori e gli innovatori possono trovare nel territorio un *"laboratorio vivente"* in cui sperimentare le loro soluzioni sui dati del territorio, verificarne l'efficacia e la sostenibilità.

Per fare questo, è necessaria una forte interazione fra pubblico e privato. La Fondazione Bruno Kessler, è un centro di ricerca che allinea i suoi piani strategici con quelli del territorio, con i decisori istituzionali che operano sul territorio, che decidono le politiche delle città, della mobilità, della scuola, della sanità e dei servizi pubblici in genere. Partendo da questi piani strategici, le attività di FBK mirano a portare innovazione sul territorio secondo un paradigma di *"ecosistema di business pubblico-privato"*. FBK è l'attore fra il pubblico e il privato che deve comprendere le esigenze del pubblico e, al tempo stesso, abilitare il business delle aziende private (Figura 1).

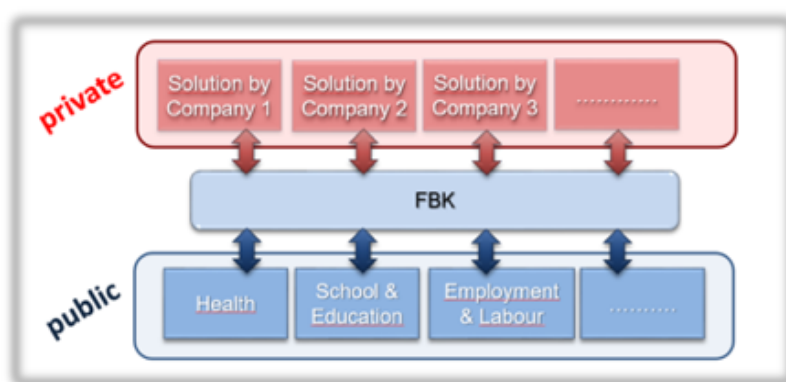


Figura 1

FBK: Un attore chiave fra pubblico e privato

FBK come attore chiave fra Pubblico e Privato, in grado di comprendere e farsi carico, per quanto riguarda il processo di digitalizzazione, dei bisogni del pubblico (della pubblica amministrazione) e al tempo stesso abilitare un business privato

Essendo FBK un centro di ricerca che produce tecnologia innovativa, ed in particolare tecnologia digitale, il ruolo di FBK diventa quello di costruire piattaforme tecnologiche che permettano di rispettare i requisiti del pubblico. Queste piattaforme al tempo stesso devono permettere ad aziende private di costruire le loro soluzioni di business, aprendo in questo modo alle aziende un mercato che rispetti i requisiti voluti dal pubblico (Figura 2). La piattaforma deve essere una piattaforma aperta, ingegnerizzata al punto da favorire l'utilizzo anche da parte di piccole e medie imprese, che non hanno la forza di investire nella costruzione di tale piattaforma ma la possono utilizzare per proporre le loro soluzioni.

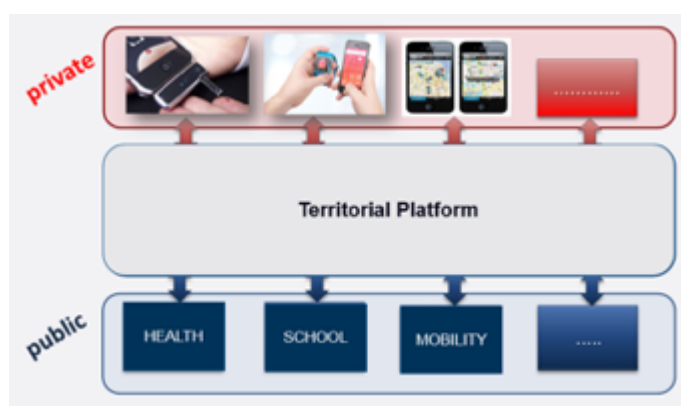


Figura 2

La Piattaforma Territoriale, un sistema digitale fra pubblico e privato

La piattaforma serve a raccogliere dati dal territorio e dalla comunità che vive sul territorio. In Figura 3 mostriamo un esempio di alcune funzionalità offerte dalla piattaforma open source, inclusi servizi chiamati "general enablers" (dedicati ad esempio alla sicurezza, ai pagamenti, etc.) e sistemi che permettono di interfacciare applicazioni di privati nel campo della sanità, della scuola, e del lavoro, mantenendo dati e servizi secondo una strategia dettata dal pubblico.

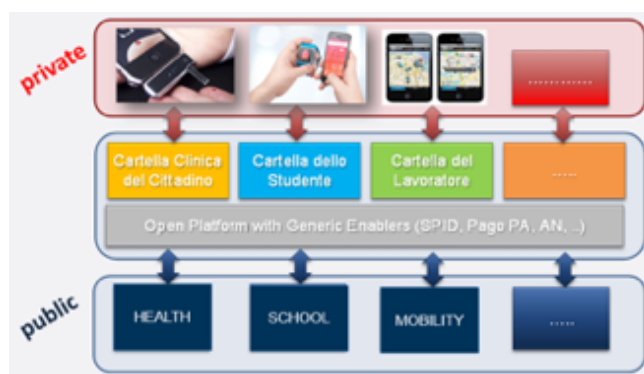


Figura 3

Un esempio di alcune funzionalità svolte dalla piattaforma pubblica abilitante il business privato

Vediamo adesso più nel dettaglio l'architettura concettuale di tale piattaforma digitale, evidenziando alcuni componenti in particolare (Figura 4). Questa piattaforma realizza la visione di una tecnologia che copre tutta la "filiera" dalle infrastrutture e i sensori, dall'internet delle cose, alla elaborazione dei dati, ai servizi digitali fino al coinvolgimento attivo delle persone. In Figura 4 partiamo dal livello più basso, l'infrastruttura, con diversi dispositivi per raccogliere dati dal territorio, l'Internet delle Cose (Internet of Things - IoT). Su questo livello poggiano le strutture per raccogliere e gestire i dati e fornire servizi. I dati territoriali sono aperti ("*Open Data*") in modo che siano disponibili a chiunque li voglia utilizzare, ed in particolare anche alle aziende private. Un certo tipo di dati sensibili, invece possono essere raccolti in un "*Personal Data Storage*". Sui dati si costruiscono servizi digitali aperti ("*open service*") in modo che aziende private possano costruire i loro servizi digitali utilizzando questi servizi come base. Si arriva poi alla parte superiore che è dedicata al coinvolgimento attivo della popolazione. Tramite ad esempio tecniche di "*gamification*" che coinvolgono con giochi le persone stimolandole ad adottare ad esempio sani stili di vita. Infine tecniche di "*idea management*" permettono ai cittadini di proporre idee per servizi innovativi che possono essere poi approvati e realizzati dai decisori istituzionali.

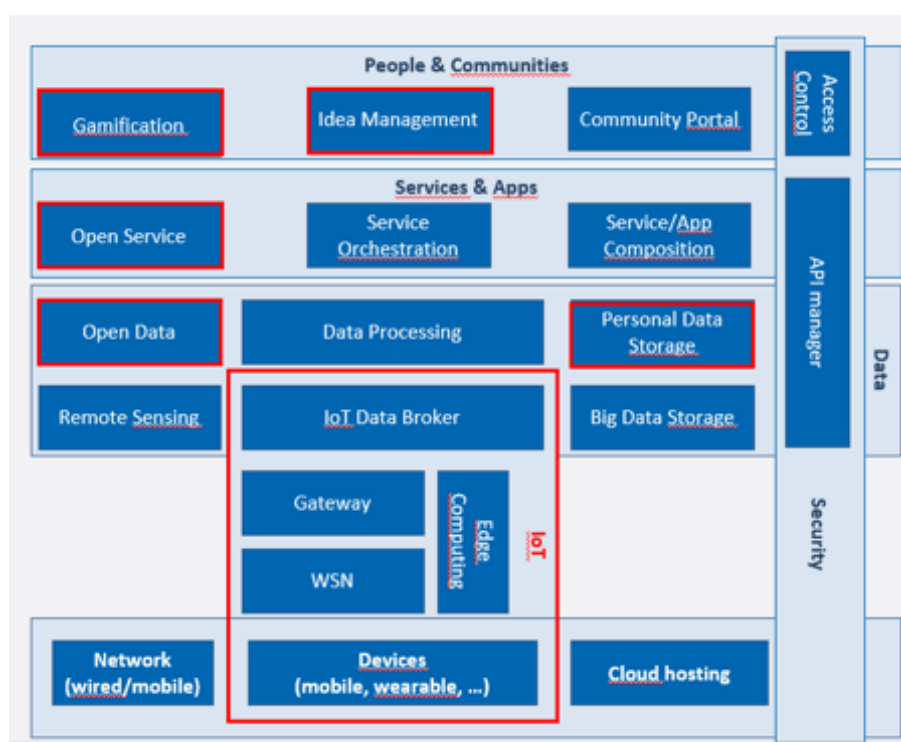


Figura 4
Architettura Concettuale della Piattaforma Territoriale

Sono questi i componenti alla base di una tecnologia che soddisfa i requisiti voluti dal pubblico, dai decisori istituzionali e al tempo stesso in grado di

abilitare un business privato. Uno degli esempi più significativi di *"Personal Data Storage"* è la *"Cartella Clinica del Cittadino"* che permette alle persone di raccogliere dati personali sulla loro salute. Questi possono essere poi utilizzati da diverse applicazioni, ad esempio per il monitoraggio di malattie croniche, o da applicazioni per la promozione di sani stili di vita. Questo è un chiaro esempio di soluzione tecnologica che soddisfa un requisito voluto dal decisore pubblico ma che permette a privati di costruire applicazioni che soddisfino i requisiti di privacy e possano essere soluzioni di mercato. La cartella clinica del cittadino è oggi utilizzata da più di 80.000 persone e apre la strada a sistemi di *"virtual coaching"*, ovvero agenti conversazionali che interagiscono con le persone per guidarle, anche fuori dalle strutture ospedaliere, nella gestione della loro malattia o che aiutino la persona ad adottare sani stili di vita. Il sistema di *"virtual coaching"* basato sulla cartella clinica del cittadino, diventa così un *"digital twin"* della persona per quanto riguarda la sua salute, un chiaro esempio a mio parere di *"sistema digitale simbiotico per la persona"*. Ma un *"digital twin"* della persona non si può occupare solo di salute, ma anche di tutti gli aspetti della vita della persona. Per questo stiamo costruendo in Trentino, sempre con lo stesso approccio, i sistemi che aiutino gli studenti a gestire lo studio e l'alternanza scuola lavoro, i lavoratori, le persone anziane, in genere i cittadini, in modo da estendere il concetto da *"cartella clinica del cittadino"* e relativo virtual coaching per la salute a *"cartella della persona"* e relativo *"virtual coaching"*.

Per rendere più concreta questa visione, nel seguito descriviamo brevemente un progetto della Fondazione Bruno Kessler (FBK) che si è ispirato alle idee che ho appena illustrato. Il progetto *"Children Independent Mobility"* (CLIMB) utilizza tecnologie digitali avanzate per la mobilità, la formazione e per stimolare la vita sociale dei bambini.

CLIMB: Children Independent Mobility

La mobilità attiva e sostenibile dei bambini è fondamentale per lo sviluppo fisico, sociale, cognitivo ed emozionale del bambino e contribuisce a definire il rapporto del bambino con la città. Attraverso il progetto CLIMB (<http://climb.smartcommunitylab.it>), la Fondazione Bruno Kessler sta affrontando la sfida di utilizzare ed integrare tecnologie informatiche avanzate e di metterle a servizio della comunità per una gestione più efficace e dinamica della mobilità dei bambini.

Il progetto CLIMB è iniziato come attività all'interno di un laboratorio sperimentale in un quartiere nella città di Trento (un *"Living Lab"*), il rione la Vela, nato dalla collaborazione tra Comune di Trento e FBK, il cui obiettivo è introdurre e testare soluzioni digitali innovative e prevede: l'installazione di strumentazione e infrastrutture ICT (es. sensori intelligenti, apparati di comunicazione innovativi); l'analisi e al monitoraggio dello stato del quartiere (*"city sensing"*); la creazione di cruscotti informativi dedicati alla pubblica amministrazione e alla cittadinanza; la realizzazione di soluzioni informatiche (es. app, portali, strumenti di collaborazione e partecipazione) per i cittadini; il coinvolgimento della comunità locale nei progetti di innovazione e sperimentazione.



Figura 5.
Il Laboratorio Sperimentale nel Quartiere Vela di Trento

Nel progetto CLIMB la tecnologia sostiene la riconquista della mobilità indipendente attraverso due soluzioni tecnologiche: *"Pedibus Smart"* e *"Kids Go Green"*. *"Pedibus Smart"* è una soluzione, supportata da App mobile, dedicata ai volontari del pedibus e mirata a semplificarne le attività di gestione (organizzazione turni dei volontari, sostituzioni, gestione dei bambini/passeggeri e presenze giornaliere, contatti con i genitori). L'App è corredata da dispositivi smart, che sono distribuiti ai bambini, e che, sfruttando sensori di prossimità, permettono di automatizzare la registrazione della presenza alle diverse fermate e la compilazione del "diario di bordo". *"Kids Go Green"* è un sistema di gioco che vuole promuovere la mobilità sostenibile e attiva dei bambini tramite la tecnologia persuasiva della *"gamification"*. In questo gioco, sommando tutti i Km sostenibili (a piedi, in bici, con lo scuolabus) fatti giornalmente dai bambini, si fa avanzare la scuola in un cammino nel mondo reale. Cammino e tappe sono personalizzabili in base agli interessi e al programma scolastico della scuola / classe. Inoltre, ogni classe può associare del materiale didattico multimediale di approfondimento alle diverse tappe. La App del *"Pedibus 2.0"* permette di automatizzare il gioco, raccogliendo in modo automatico i Km fatti da bambini e volontari del pedibus.

Le soluzioni CLIMB sono utilizzate in diverse scuole del Comune di Trento e del Comune di Ferrara. Durante l'anno scolastico in corso (2018-2019) sono state attivate 9 linee *"Pedibus Smart"* e 19 percorsi *Kids Go Green*, per un totale di oltre 1000 bambini, 100 insegnanti e 60 volontari coinvolti in 14 scuole. Le soluzioni CLIMB sono utilizzate da tre anni in diverse scuole del Comune di Trento. Durante l'anno scolastico 2017-2018 sono stati attivati 3 *"Pedibus Smart"* e 10 giochi *"Kids Go Green"* nelle scuole primarie, per un totale di oltre 600 bambini e 100 insegnanti coinvolti. Queste sperimentazioni hanno un impatto sulla mobilità casa-scuola che ha superato ogni aspettativa: nelle scuole che hanno sperimentato CLIMB quest'anno, i bambini hanno percorso più di 20.000

Km sostenibili; e dei 26.000 tragitti casa-scuola, solo il 6% è stato fatto in auto. Le sperimentazioni hanno mostrato un impatto sulla mobilità casa-scuola che ha superato ogni aspettativa: nelle scuole che hanno sperimentato CLIMB nell'a.s. 2017-2018, i bambini hanno percorso più di 20.000 Km sostenibili; e dei 26.000 tragitti casa-scuola, solo il 6% è stato fatto in auto. I volontari del piedibus riportano che la soluzione Piedibus Smart ha semplificato la gestione giornaliera del servizio piedibus e permette loro di concentrarsi sugli aspetti di sicurezza e di relazione con i bambini.



Figura 6

La soluzione Pedibus Smart utilizza tecnologie IoT per rendere la mobilità dei bambini sicura e indipendente

Il progetto ha inoltre dimostrato un alto valore educativo e didattico, soprattutto grazie ai percorsi virtuali personalizzabili e all'utilizzo di tecnologie che supportano una partecipazione e un coinvolgimento attivo degli alunni nell'apprendimento. Le insegnanti testimoniano che CLIMB è stato per loro "un modo completamente nuovo di fare didattica", che ha abilitato percorsi interdisciplinari, collaborazioni intra e inter-classe, e una partecipazione entusiastica di docenti e alunni. Le famiglie inoltre riscontrano un impatto significativo di CLIMB sia in termini di consapevolezza del bambino verso tematiche di sostenibilità ambientale, che in termini di autonomia, sia negli spostamenti casa-scuola che nel tempo libero.

Grazie al successo di queste sperimentazioni, e alla promozione dal basso del progetto da parte di famiglie e insegnanti, sono numerose le comunità scolastiche interessate ad adottare le soluzioni CLIMB. Nel Comune di Trento, tutte le scuole coinvolte hanno confermato il loro interesse a continuare o rafforzare l'adozione di CLIMB e nuove scuole hanno aderito. Prevediamo quindi significativo aumento di partecipazione sia in termini di insegnanti/alunni coinvolti che in termini di durata delle sperimentazioni. In Italia, il dialogo è già attivo con diverse città, interessate a portare questa esperienza positiva nelle loro comunità scolastiche. Un esempio è la città di Ferrara, dove è già stata

effettuata una prima sperimentazione del Pedibus Smart e una adozione più importante inizierà in autunno.

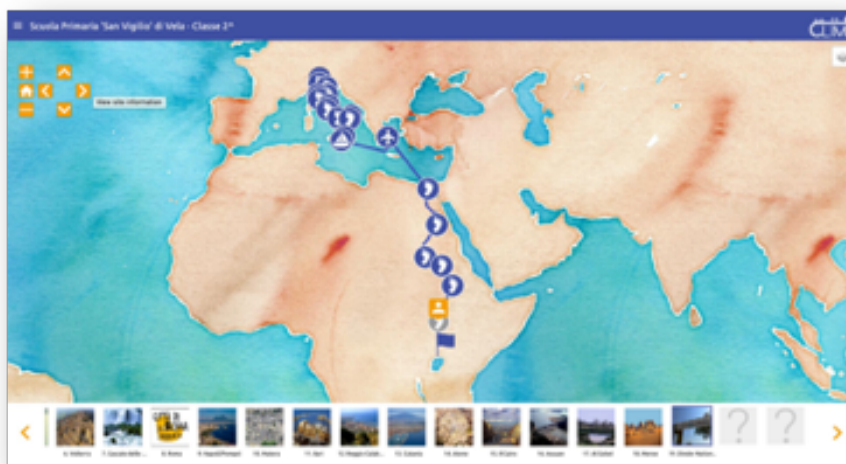


Figura 7

La Soluzione “Kids Go Green” è un sistema di gioco che vuole promuovere la mobilità sostenibile e attiva dei bambini tramite la tecnologia persuasiva della “gamification”.

Sono diversi i riconoscimenti ricevuti da questo progetto. Nel 2018 CLIMB ha vinto il premio "Wind-TRE Viabilità evoluta" nell'ambito dei CRESCO Awards. Nel maggio 2018 CLIMB è stato selezionato, nel contesto del “Premio PA sostenibile - 100 progetti per raggiungere gli obiettivi dell’Agenda 2030”, come esempio nazionale di successo in ambito mobilità sostenibile. Nel gennaio 2018 CLIMB ha vinto la Climate KIC Urban Challenge di Ferrara: un’iniziativa organizzata da EIT Climate-KIC con l’obiettivo di supportare la transizione delle città italiane verso un futuro a basse emissioni. CLIMB è inoltre una delle sette buone pratiche, selezionate tra oltre 400 proposte, che ha rappresentato l’Italia in occasione del “G7 Trasporti” svoltosi a Cagliari il 20-21 luglio 2017.

Nell’ambito del progetto CLIMB, FBK ha curato l’ideazione e la realizzazione delle soluzioni tecnologiche che ne supportano l’esecuzione. Il Comune di Trento e il Comune di Ferrara hanno promosso l’iniziativa nelle comunità scolastiche del territorio, integrandola nell’offerta formativa per le scuole primarie. Gli Istituti Comprensivi e le Scuole hanno aderito come “adopter” della soluzione, diventandone poi “champion” e promotori. Per supportare FBK negli aspetti di gestione operativa e di innovazione educativa, sono in corso collaborazioni con Kaleidoscopio, importante cooperativa di servizi per la scuola operante in Trentino, e Centro Studi Antartide, centro di eccellenza accreditato della RES, la Rete regionale dell’educazione alla sostenibilità della Regione Emilia Romagna, che agisce nel campo dello studio, dell’educazione e della comunicazione ambientale e sociale.

Nell'ambito del progetto CLIMB, FBK ha curato l'ideazione e la realizzazione delle soluzioni tecnologiche, basate su IoT e Gamification, che ne supportano l'esecuzione. Il Comune di Trento, Ufficio Politiche Giovanili, ha promosso l'iniziativa nelle comunità scolastiche del territorio, integrandola nelle proposte del progetto "A piedi sicuri". Gli Istituti Comprensivi e le Scuole hanno aderito come "adopter" della soluzione, diventandone poi "champion" e promotori. Per supportare FBK negli aspetti di gestione operativa e di innovazione educativa, è inoltre in fase di definizione un accordo con Kaleidoscopio, importante cooperativa di servizi per la scuola operante in Trentino.

In conclusione, le attività in corso nel progetto CLIMB sono un esempio della visione qui illustrata verso un sistema simbiotico digitale per la persona. L'ideazione e la realizzazione di CLIMB sono dovuti ad un intenso co-working fra esperti di tecnologia (FBK), istituzioni pubbliche (Comune di Trento, Sistema Scolastico), istituzioni private (Kaleidoscopio, Centro Studi Antartide) e persone attive sul territorio, naturalmente gli insegnanti, ma anche i genitori e persino i bambini. Nonostante la soluzione tecnologica sia altamente interattiva, il livello di automazione è elevato: il sistema registra automaticamente ogni bambino che approccia il pedibus, automaticamente vengono rilevate situazioni di pericolo, come ad esempio il fatto che un bambino si allontana dal gruppo, avvertendo tempestivamente gli accompagnatori. Questo è il primo tassello, che apre le porte ad una automazione "sociale", "per la persona" molto forte. A valle del progetto e dopo vari incontri con i genitori, in FBK stiamo studiando tecnologie che permettano ai genitori, nel rispetto della privacy e della persona, di definire dei requisiti di alto livello di sicurezza per i loro bambini, non solo mentre vanno a scuola, ma durante tutto l'arco della vita quotidiana, ad esempio mentre giocano in un parco. Quale sistema simbiotico può essere migliore se non quello che permette ad un genitore di essere tranquillo mentre la sua bimba gioca con gli amici in un parco in assoluta libertà?

Biografie

Paolo Traverso è il Direttore del Centro per le Tecnologie dell'Informazione e della Comunicazione della Fondazione Bruno Kessler (FBK ICT IRST) dal 2008. Il Centro conta più di 200 dipendenti (ricercatori, sviluppatori, project managers e studenti di dottorato) che lavorano nei settori dell'intelligenza artificiale, della scienza dei dati e dell'apprendimento automatico, dei sistemi per la sicurezza.

Email: traverso@fbk.eu

Marco Pistore è ricercatore senior in Fondazione Bruno Kessler, dove dirige la linea di ricerca su Città e Comunità Intelligenti, un gruppo multidisciplinare di ricercatori e tecnologi che lavorano in diverse aree legate alla sostenibilità e alla qualità della vita nelle città e nelle aree periferiche. Marco ha svolto attività di ricerca nei campi dell'intelligenza artificiale, delle tecnologie software, dell'e-government e delle smart city; su questi temi, ha pubblicato oltre 120 articoli in riviste, conferenze e workshop internazionali ed è stato responsabile di progetti di ricerca e industriali.

Email: pistore@fbk.eu

Elisabetta Farella è ricercatrice e capo dell'unità Energy Efficient Embedded Digital Architectures (E3DA), presso la Fondazione Bruno Kessler. La sua attività di ricerca è nel campo dei sistemi embedded autonomi energeticamente e al contempo dotati di intelligenza artificiale on board. Sono esempi di tali sistemi le Wireless Sensor Networks (WSN), l'elettronica indossabile, Internet-of-Things dal punto di vista dei device dotati di sensori e attuatori intelligenti. Tali tecnologie sono utilizzate in vari ambiti applicativi, dalla riabilitazione motoria, all'interazione uomo macchina e particolarmente nell'ambito smart cities and communities.

Email: efarella@fbk.eu

Amy L. Murphy è ricercatrice senior presso l'unità di ricerca Energy Efficient Embedded Digital Architectures (E3DA) della Fondazione Bruno Kessler. La sua ricerca si concentra sulla ricerca applicata per smart cities a partire dall'ingegneria del software, dal calcolo distribuito, dalle reti wireless a basso consumo energetico, con una recente enfasi sulle Internet of Things. Il tema che guida il suo lavoro è quello di abilitare applicazioni affidabili per ambienti dinamici con particolare attenzione ai protocolli di comunicazione wireless, necessari per supportare interazioni complesse tra dispositivi distribuiti.

Email: murphy@fbk.eu

Annapaola Marconi è la ricercatrice a capo dell'unità Distributed Adaptive Systems (DAS) presso la Fondazione Bruno Kessler. I suoi interessi di ricerca includono i sistemi distribuiti adattativi, i sistemi persuasivi e motivazionali basati su tecniche di gamification e gameful design. È attivamente coinvolta nella realizzazione e divulgazione di iniziative e programmi volti ad applicare i risultati della ricerca in applicazioni concrete in grado di creare un impatto sul territorio locale e di migliorare la qualità della vita dei suoi cittadini e delle comunità.

Email: marconi@fbk.it

Pamela Forner è Project Manager del centro ICT della Fondazione Bruno Kessler di Trento. Dopo essersi occupata dell'organizzazione di campagne di valutazione nell'ambito NLP e della creazione di risorse linguistiche, nel 2013 si è unita al gruppo di Project Management. Si occupa della gestione di progetti - a livello locale, nazionale e internazionale - in ogni fase: pianificazione, realizzazione e controllo. Ha conseguito la certificazione PMP e AgilePM.

Email: forner@fbk.eu

Towards Evolving Symbiotic Education Based on Digital Twins*

Wiltold Kinsner*

Abstract

The industrial revolution has accelerated the need for many skilled individuals that could not be produced by the personalized master-student educational system of the old. What was needed was a well-organized one-fits-all educational system that could produce expected numbers of individuals with expected skills within a fixed period of time. The Prussian classroom education system of 1770 produced magnificent results. At that time, the acquired knowledge and skills lasted for at least a lifetime. The industrial revolutions since then have accelerated the pace of knowledge doubling from a lifetime to months. Are we capable of adjusting to that pace? Furthermore, since the resilience of jobs has also been below a single lifetime, young professionals are expected to have more than one job. How can they learn all of that in the old educational system? The time has come to revamp the educational system at the core. The new system must be personalized to match the diversity of individual abilities and styles of learning. The new system must also be based not only on the body of knowledge (BoK), but body of experience (BoX). We envisage that the new personalized system of education being sufficiently agile and interactive so that it would become evolving in its symbiosis with humans. For that to happen, we must coexist with symbiotic autonomous systems, specifically involving digital twins. This paper addresses some aspects of this view.

Sommario

La rivoluzione industriale ha aumentato la necessità di disporre di molte persone con uno skill elevato che non potevano essere "prodotte" attraverso un insegnamento personalizzato quale quello utilizzato precedentemente. Diventava necessario passare ad un

*Those interested in getting a more in depth vision could refer to the documents available in the Symbiotic Autonomous Systems Initiative Education website: <https://symbiotic-autonomous-systems.ieee.org/education>

sistema educativo in grado di gestire elevati volumi in tempi contenuti con un approccio “un solo sistema adatto per tutti”. Il sistema educativo Prussiano del 1770 basato sulla educazione in classe portò a risultati eccellenti. A quell'epoca le conoscenze acquisite duravano per tutta la vita. La rivoluzione industriale ha accelerato la crescita di conoscenze diminuendo al contempo la sua durata, da una vita a qualche anno/mese. Siamo in grado di adeguare il nostro sistema educativo a questa nuova situazione? Inoltre anche la durata dell'impiego non è più coincidente con la durata della vita lavorativa, i giovani devono aspettarsi diversi mestieri nel corso della loro vita. Come possono imparare questo con sistemi educativi pensati per rispondere ad esigenze diverse? È arrivato il momento di ripensare il sistema educativo dalle sue radici. Il nuovo sistema dovrà essere personalizzabile alle diverse capacità e necessità di apprendimento. Deve essere basato non solo su di un Body of Knowledge (BoK) ma anche su di un Body of Experiences (BoX). Occorre pensare ad un sistema educativo in simbiosi con le persone e l'utilizzo di tecnologie e paradigmi come i Digital Twin possano realizzarlo. Questo articolo riflette su questi aspetti.

Keywords: Symbiotic education; digital twins; symbiosis; body of knowledge - BoK; body of experience -BoX; life-long learning; learning ecosystem; knowledge pyramid; closed-loop education; symbiosis self-organization

Digital Twins

Digital Twins in Industry

A number of industries like General Electric (GE), Tesla, NASA, are creating digital twins defined as digital replicas of their products like airplanes, cars, and satellites. The idea is to mirror a physical analog object in bits (i.e., a physical digital system, not resembling the original object in shape, but in its behaviour) keeping the bit replica synchronized with the physical one. This allows various types of retrospective and predictive analyses on the digital twin that can provide a better insight into the analog one, and lead to corrective actions, when required. In this sense, digital twins are new tools for education: rather than studying and training on the analog object, one can study on its digital representation first. Many technologies like virtual reality can further enhance training and education.

The usefulness of a digital twin goes beyond that scenario. The digital twin can develop far beyond our physical and physiological limitations, and find proper ways to be helpful in our adaptation to the untenable challenge of doubling of knowledge over a decreasing time period. Another challenge is the need to educate individuals for more than one job due to automation, mechanization,

and the unprecedented growth in deep learning (DL) [1] and artificial general intelligence (AGI), as described in the White Paper I and II. Some challenges in developing better engineering education in cognitive systems are described in [2], [3].

Is There a Digital Twin of Me Already?

In a way, each of us has already several fragments of our own digital twin. Social media like Instagram and Facebook, LinkedIn and Twitter are collecting parts of our “self”. Governments and municipalities are also collectors of parts of our “self.” The health-care system “knows” much about our body and mind. Large physical department stores, as well as digital merchandise systems like Amazon “know” much about our purchasing needs and interests (electronic products, mechanical gadgets, books, music) so that they often suggest new and related products. Google has a deep insight into its users’ interests ranging from scientific, technical, conceptual, philosophical, artistic, political, social, to theological. Travel companies know our interests about the world and our disposable income. In addition, the companies where we have been working and where we work now have other fragments, representing our acquired skills and habits. Insurance companies know much about our health risks. The educational institutions that we have used are also collecting records of what we have learnt, how good we are at specific subjects, and can assess our intellectual potential. All of those fragments approximate of what and who we are. They are distributed elements of our digital twins.

At this stage, all those fragments are dispersed. Some countries are developing rules to establish ownership of those fragments. For example, Italians have the right to access these data and information, and the companies physically storing the data have to grant them access. Having the right and actually being able to access them easily, are quite different stories.

Digital Twin: Aggregation of Fragments

In perspective, we should be able to aggregate those fragments into a more comprehensive one in order to represent our “self” better. In addition, it is most likely that the number of collected observations, data, and extracted information about ourselves will grow in time, thus leading to more and more accurate representation of our “self”.

If we imagine a symbiotic relationship between a person and the corresponding digital twin, the symbiotic counterpart could form a very good understanding of who we are, sometimes through direct access to what we do jointly, some other times through the access to other digital twins.

A Symbiotic Digital Twin

In a formal way, our digital twin could come to represent both our skills, knowledge, and wisdom. It can also be flanked by applications taking into account the fading away of skills (what we lose when not practising) and knowledge (when we forget). This information of our degrading skills/knowledge can be the starting point for a proactive education program.

Writing an article and presenting it at a conference, or attending a conference to listen to colleagues presenting their papers can also be mirrored by our digital twin. The same applies to the process of reviewing papers. Many publishers allow ongoing discussion on their published papers that could be monitored by our digital twin. Educational institutions, including IEEE, could contribute to the mirroring of their “students”, “members” into digital twins. These might come handy in creating customized and personalized education programs. An example of such a program is the *personalized system of instruction* (PSI) by Fred S. Keller (1899-1996; 97) [1]. Since the manual administration of Keller’s PSI is very tedious, we have developed a *Computer-Aided PSI* (CAPSI) that has been running at the University of Manitoba, Canada for many years [5].

In a symbiotic autonomous system (SAS), the skills, knowledge and wisdom should be shared among its component subsystems to enhance the overall performance of the system. Furthermore, the digital twin could start increasing (or decreasing) interaction between its component parts. Notice that in dynamical complex systems, the whole is not necessarily the sum of its part. Through such nonlinear interactions, an emergent quality may appear that may not be found in any of its parts.

How Can a Symbiotic Digital Twin Help Me in Skill and Concept Learning?

If I live in a symbiotic relation with my appliances at home or at work, the knowledge of what specific selection/action/effect (a “program” for short) I would most likely to be interested in at a give time becomes part of the global knowledge of the symbiotic digital twin. However, the knowledge about what programs are available and would fit my interest may lie in an appliance. Notice that Amazon’s Alexa, Apple’s Siri, Microsoft’s Cortana are all moving in this direction. There are now thousands of programs (such as streaming contents) to choose from and be of serious interest to me, but they are just too many for me to be aware of at any given time. The same applies to the millions of YouTube clips, tweets, and other pieces of information that could become an integral part of my education process, but I will never know that they even exist.

The same scenario can be envisaged with studying magazines, white papers, reports, textbooks, monographs, and research papers. Another scenario emerges with all the courses and MOOCs (massive open online courses) that are available at Coursera launched in 2012 by Stanford University (over 2,000 courses), EdX launched in 2012 by Harvard University (over 1,200), Udemy (over 2,500), Udacity (200), MIT OpenCourseWare (2,200), XuetangX founded in 2013 by Tsinghua University (over 500), Lynda (3,300), Khan Academy started in 2006 by Salman Khan, TED (1,890), The Great Courses (500), and over 700 universities offering MOOCs (e.g. [6] [7]). Designing of online courses is discussed in [8] and improved comprehension in [9].

A similar scenario emerges when doing research. Finding and reading relevant research papers, technical magazines, technical reports, white papers, technical books is very time consuming. These are all examples of the “where can I find it?” problem in education. A digital twin could help in these situations.

More Reasons for Symbiotic Digital Twins: Knowledge Doubling and Its Half-Time

With the explosion of data, information, knowledge and wisdom, we would have to spend all our available time searching for what is needed for our education and work. We cannot just ask a teacher or professor to answer our questions outside their class or research area. Today, search engines still provide millions of hits that have to be reviewed for relevance. Finding the relevance in the sifted out and even prioritized material takes time. Since our reading and comprehension abilities are slow (the average reading speed is around 300 words per minute), it might take up to four hours to keep up with daily emails, news digests, blogs, magazines, and books. This keep-up time reduces the time for creative work.

According to Buckminster Fuller's "knowledge doubling curve" in 1982, all human knowledge generated and transmitted doubled in size around year 1500. It doubled again by 1750 (only 250 years), and doubled again by 1900 (just 150 years). With those rates, humans were able to adapt to the growth and change. It became harder to adapt when the doubling took 25 years around 1950. The knowledge doubling today is much shorter (around 13 months). As an example, the number of annual patents increased from about 50,000 to more than 325,000 over the last 50 years. Many in IBM predict that in not-too-distant future (2020), the knowledge doubling will happen in 12 hours. It is not feasible for a human to adapt to that rate. The concept and implementation of a digital twin seems to be a necessity now [10].

There is another reason for digital twins: the knowledge half-life. In his book Future Shock [11], Alvin Toffler stated that “the illiterate of the 21st century will not be those who cannot read and write, but those who cannot learn, unlearn, and relearn.” The knowledge and skills acquired in our schools and successive jobs diminishes in value, and requires continuous updating, not once but throughout our lives.

How long does it take for knowledge to become outdated and irrelevant, or even incorrect? The half-life of knowledge (i.e., the amount of time it takes for knowledge to lose half its value) is often used to indicate the devaluation of knowledge in various disciplines. As might be expected, the knowledge half-life in aggressive disciplines like science, engineering and technology is shrinking fast.

Knowledge Tsunami and Organizations

The conditions when knowledge-doubling occurs exponentially, while the knowledge half-time decreases may have a tsunami effect on any society, organization, company or other organizational unit. The SAS with digital twins could be very helpful in increasing our resilience in some of the following areas: (i) Curation of knowledge (organizing and filtering according to agreed-upon criteria to eliminate irrelevant knowledge); (ii) Knowledge fusion (to discover and clean errors present in sources, as well as mistakes made in the process of knowledge extraction from sources); (iii) Plagiarism management (to generate new knowledge); (iv) Knowledge vetting (to identify and verify sources for quality of the content used in the organization); (v) Intellectual property management

(separating intellectual property, trade secrets, and copyrighted information from generic and public-domain content); (vi) Knowledge sunseting (to identify knowledge that cannot be used any longer); (vii) As traditional libraries dwindle, creation of a digital twin "librarian" that knows the needs of the organization and its members would be beneficial; (viii) As traditional publishing also dwindles, creation of a suitable digital twin "publisher" could benefit the organization.

From Competition to Cooperation and Knowledge-Sharing Society

In the past, individuals and teams in companies and organizations worked independently, competing in the quest for reaching the finish line first. Competition has been recognized as healthy, but leaves many in the dark.

The emerging possibility of educational digital twins in the SAS environment creates a set of new opportunities for knowledge to be uncovered and discovered, distributed widely, and then enhanced by the same individuals, but now working in the SAS environment. In fact, new knowledge is now developed mostly by collaborating or even cooperating teams rather than individuals.

What Prevents Us from Having the SAS in Education Today?

There are at least three missing links to create a SAS in education: (i) a reasonable digital twin itself, (ii) symbiotic interconnects, and (iii) rigorous modelling of SAS.

While the concept of a digital twin is not new, the concept of a digital twin of a human is fairly new. While attempts are being made to develop machines and systems capable of acting in a manner that could be described as: adaptive, autonomous, intelligent, perceptual, cognitive, conscious, and symbiotic with humans and the environment, digital twins are not here yet.

The second missing link is the connectivity between the different parts of the symbiotic system. When the Internet has been developed sufficiently, we thought, happily but mistakenly, that education could be improved on a dime by delivering many online courses (in different forms such as the massive open online courses, MOOCs) to many people at any time and any place. Of course, this helped, as millions of new individuals obtained access to education that was not available to them before. However, the interconnections were not symbiotic. We must develop a new class of symbiotic interconnects.

The third missing link is proper rigorous modelling of SAS based on brain-inspired and socially-inspired processes. This modelling also implies quintessential changes in signal processing in the simulation of the processes. In the past, the majority of signal processing was done on a single scale (mono-scale). More advanced models of reality included multi-scale signal processing. Cognitive system and SAS require not only more elaborate multi-scale, but also poly-scale modelling and signal processing. Some of the definitions will be provided in the next subsection, and summaries of the modelling techniques may be described in SAS White Paper III.

How Can We Get There?

The missing links can be addressed by some innovators, including some specific initiatives by IEEE or ACM or other organizations (obviously not related

to the above case on how to learn using a photographic camera to make aesthetic photographs in a reasonably short time, but, as an example, to support careers path of its IEEE members).

Towards Evolving Symbiotic Education

Summary of Reasons

As discussed already, the knowledge tsunami and automation has put much pressure to change the current classroom/workshop model in vocational training. More students and workers learn “just-in-time”, and often just enough to solve a problem or get a job completed. Teachers and trainers can no longer be the main sources of knowledge about the world of work, but need new forms of technology to help find and manage the increasing amount of information. No single person, no matter how brilliant, can handle the knowledge, even in one field of study.

Consequently, the roles of teachers, trainers and consultants need to change – from mostly presenters of information to guides, mentors, curators of knowledge, critical thinkers, and problem solvers. They will have to use digital learning skills and literacies.

I have been making the case that the next generation of education would benefit much from the development of symbiotic digital twins capable of being in relation with human beings in a symbiotic system. This intimate knowledge of the personal needs and abilities could allow the digital twins to deliver both the BoK and BoX in a personalized way that has a chance to compete with the best model of the Oxford Master of the past.

Model of a Digital Twin in Education

A model of the digital twin that could be used in the symbiotic education is illustrated in Fig. 1.

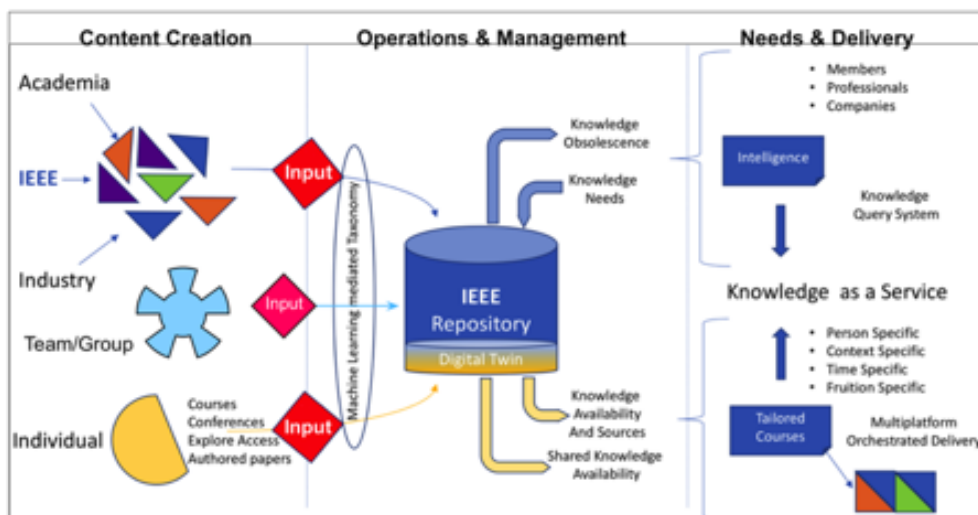


Figure 1
A digital twin model for symbiotic education.

For clarity, the digital twin is segmented into three parts: (A) content creation, (B) operations and management, and (C) needs and delivery. Each part is also subdivided into several groups of activities. For example, Part A has three distinct important content creators: (1) organizations, (2) groups, and (3) individual.

A. Content Creation

The first content creator (A1) includes: (i) the academic research and education, (ii) industrial research and development, together with the cumulative experience in the corresponding field, (iii) business modelling and analysis, and (iv) technical and scientific organizations such as the Institute of Electrical and Electronics Engineers (IEEE), the Institution of Engineering and Technology (IET, formerly the Institution of Electrical Engineers, IEE), the Association of Computing Machinery (ACM).

The second content creator (A2) includes (i) teams and (ii) groups of individual whose objectives and driving forces may be fundamentally different from the corporations and organizations of Class A1. Their impact may also be very different in scope. Disruptive technologies and ideas often start at that level.

The third distinct content creator (A3) is the individual who operates differently from the other two content creators. Individuals do research, write papers, read papers, access IEEE Xplore, connect dots or relations not seen by others before, create courses, contribute to conferences, contribute to teams, as well as contribute to corporations, countries, and societies.

What is not shown in Fig. 5 is a content creator that is larger than corporations and organizations: the society and its culture, together with both the environment that the society exists and its humanity. Without closing this loop between the individual and humanity, the process may become unsustainable and meaningless. But it has never been so, in the long run.

B. Operations and Management

This operations and management part of the digital twin model includes three key constituents: (i) a repository of all that is known, (ii) a master analyzer and taxonomy creator (MaX), based on the inputs from Part A and from the knowledge repository, and (iii) the part of the digital twin that is visible to the outside world.

Based on the current needs, the digital twin provides the BoK and BoX that is available and suitable to the individual recipient. Using the best available machine learning mechanisms, the MaX categorizes the enormous influx of new knowledge and experience, and evaluates obsolescence and possible sunseting of parts in the current BoK and BoX.

The MaX concept has the potential to induce new approaches, methods and techniques to learn by machines to help learning by humans. Why deep learning has improved the level of sophistication in machine learning, much more has to be considered from the perspective of an individual digital twin and the society of digital twins.

C. Needs and Delivery

This part of the system is responsible for identification of individual needs and personalized methods of delivery. The needs may be related to individuals, professionals, practitioners, members of an organization, companies, businesses, governments, countries, societies. The needs may stem from either personal or team interests and may be identified based on the specific (a) context, (b) time and (c) expected outcomes or any combination of the three.

There is another possible outcome of using digital twins. Business intelligence gathering in cyberspace from individuals, companies, organizations, governments, countries and economic systems is already practiced today. Unchecked and unethical practices may, however, lead to a collapse of the system itself, as individuals may select to see the profitable part of the system only. We see an opportunity for the digital twins to help in resolving this age-old problem. Digital twins may actually enforce the ethical use of the system by operating ethically.

Expected Outcomes from Using Digital Twins: Symbions

The result of such an evolving symbiotic educational system is illustrated in Figure 2. This figure shows the digital twin as an inverted pyramid. This is to emphasize the **complementary symbiotic relation** between an individual and the DT system. Such an intertwined DT and a person could be called a **symbion**. The diagram is also intended to imply that the symbions could penetrate the environment even deeper than the human alone. The system could participate in the data mining and processing to extract information more relevant to the individual and to the environment. It could possibly see more patterns in the information and extract more significant knowledge that could be used in the decision-making process. This layer is now the widest because the digital twin would be connected symbiotically to all the relevant other digital twins for consultation.

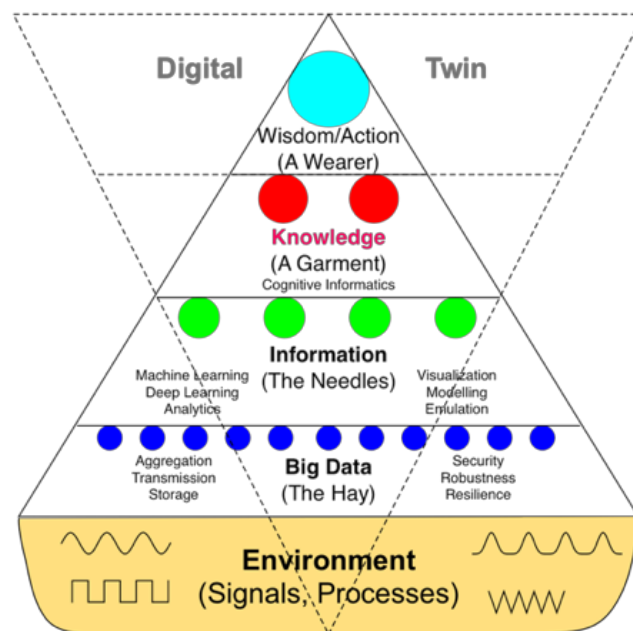


Figure 2
The impact of symbiotic education: Symbions.

Possible Outcomes from Symbiosis: Deeper Insight and Purpose

Learning through the symbiotic relation between an individual and a digital twin is potentially much more beneficial to the individual than alternative learning methods. The symbiotic pair (the individual and the digital twin) is called here a symbion. It should be very clear that such individual symbions are connected with other symbions through the fundamental construct of a digital twin. The individual symbions can enter into symbiotic relations with other symbions, forming teams, communities, and societies, as shown in Fig. 3b.

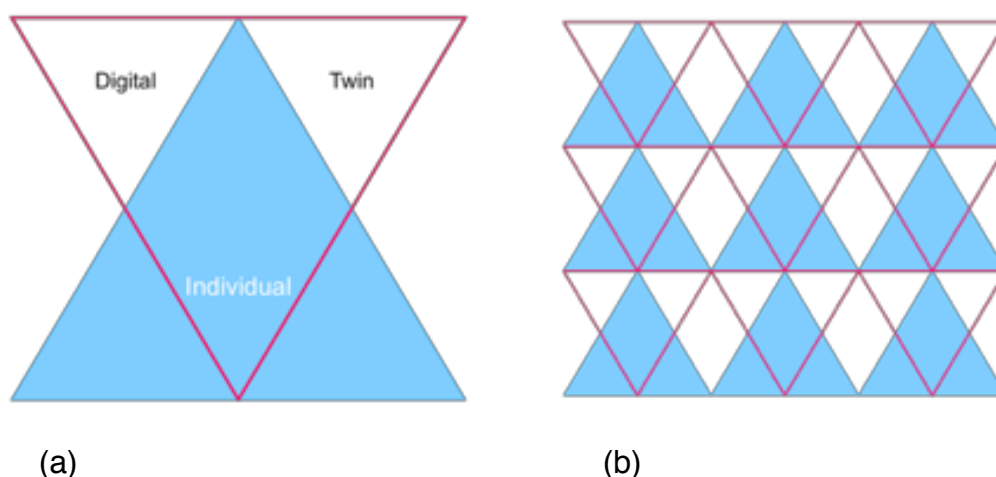


Figure 3

The impact of symbions: New societies and new knowledge. (a) A single symbion. (b) A small self-affine team.

The team shown in Fig. 3b is uniformly symmetrical. This pattern can be resized and repeated at different scales. It forms a self-similar object. The symbions can be displaced, flipped and rotated in either a deterministic or a random fashion, thus forming a self-affine structure, as shown in Fig. 4.

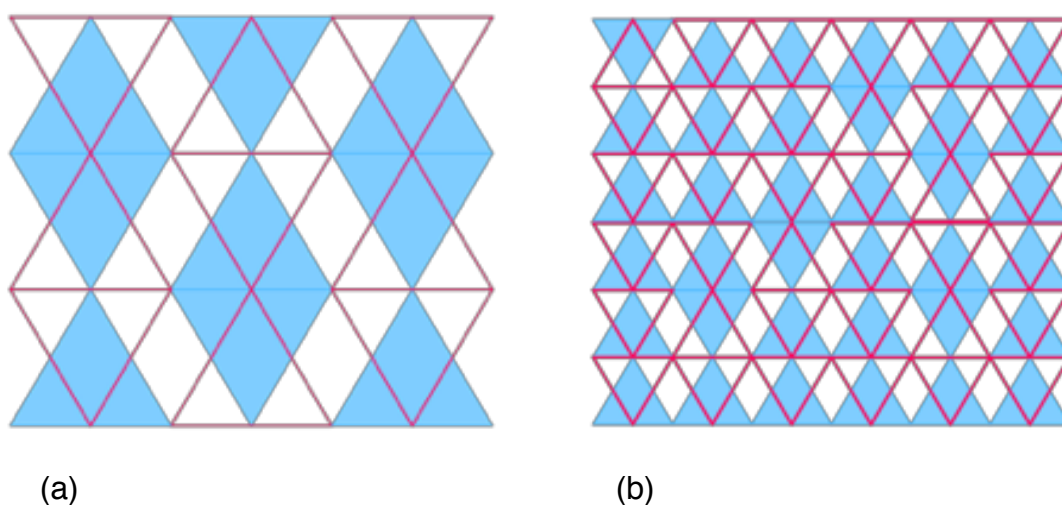


Figure 4

Symbions groupings: (a) Self affine, (b) Stochastic self affine.

The structures can be further altered to achieve even different topologies, as shown in Fig. 5.

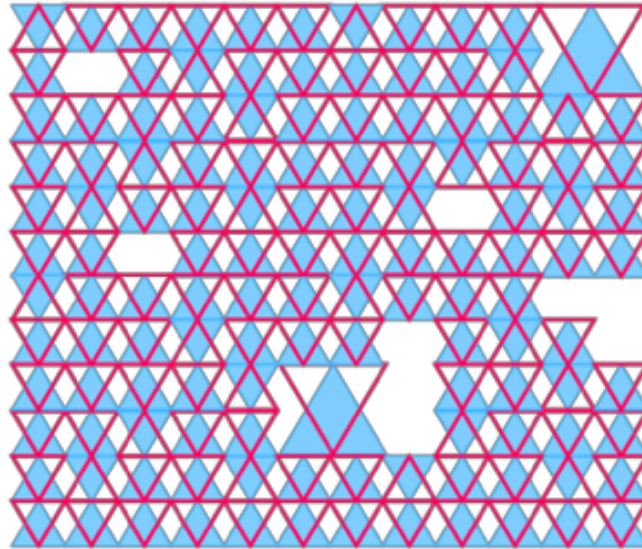


Figure 5
Stochastic topological groupings.

The advantage of this modelling is the ability study the dynamic behaviour of such teams and societies, using multiscale and polyscale techniques capable of identifying long-range dependencies, that are critical in predictive control in cognitive systems.

Closing Remarks on Symbiotic Education

- Symbiotic education has the promise of great impact on how we study, learn, acquire skills, interact with people and machines, discover new things, learn how to operate new things, and how to see reality much deeper.
- Symbiotic education can open up a new landscape for exciting new concepts and research projects.
- We already know how to compete. Symbiotic education might help us how to compete fairly.
- While competition could improve in fairness , we might also learn how to cooperate better.
- To succeed, symbiotic education must use the most sophisticated algorithms available today, and might accelerate development of better algorithm including:
 - Deep learning and machine learning [12];
 - Cognitive systems [13], [14];

- Web intelligence [15];
- HigherOrder (HO) statistical signal processing [16], [16];
- Intelligent signal processing [54, [18]];
- Compressive sensing [19];
- Fuzzy and granular computing [20];
- Multiscale (wavelet) analysis [21], [30];
- Polyscale and fractal analysis [58], [30];
- Long-range-dependence patterns in the data [22], [23];
- Nonlinear time series analysis [59];
- Emergent dynamical systems concepts [24], [25].

References

[1] Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning. Cambridge, MA: MIT Press, 2016, 800 pages. {ISBN: 978-026203561-3}

<https://mitpress.mit.edu/books/deep-learning>

HTML form:

<http://www.deeplearningbook.org/>

[2] Witold Kinsner, "Challenges in the design of adaptive, intelligent and cognitive systems," Intern. J. Software Science & Computational Intelligence, vol. 1, no. 3, pp. 16-35, July-Sep 2009.

[3] KHWP12] Witold Kinsner, Simon Haykin, Yingxu Wang, Witold Pedrycz, Ivo Bukovsky, Bernard Widrow, Andrzej Skowron, Piotr Wasilewski, and Menahem Friedman, "Challenges in engineering education of cognitive dynamic systems," in Proc. of the Canadian Engineering Education Association Conference, CEEA 2012, (Winnipeg, MB, Canada; June 17–20, 2012), Paper 119, pages. 51-62, 2012. Accessed Mar 2015 from CEEA at

<http://library.queensu.ca/ojs/index.php/PCEEA/article/viewFile/4633/4615>

<http://www.ceeaa.ca/images/content/ceeaa12-proc-complete-v35s.pdf>

[4] Fred S. Keller, "Goodbye teacher...", Journal of Applied Behavior Analysis, vol 1, pp. 79-89, Spring 1968. Available Aug 15, 2018 from NIH at

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1310979/pdf/jaba00083-0078.pdf>

[5] Witold Kinsner and Joseph J. Pear, "Computer-aided Personalized System of Instruction for the virtual classroom," Can. J. Educational Communication, vol. 17, no. 1, pp. 21-36, 1988.

[6] Antonio Fini, The Technological Dimension of a Massive Open Online Course: The Case of the CCK08 Course Tools," International Review of Research in Open and Distributed Learning, vol. 10, no. 5, 2009.

[7] G2Crowd, "105 Best Online Course Providers 2018," G2 Crowd, 2018.

<https://www.g2crowd.com/categories/online-course-providers>

[8] Terry Anderson (ed.), The Theory and Practice of Online Learning. Edmonton, AB: AU Press, 2008 (2nd ed.), 484 pages. {ISBN: 9781897425077, PDF, free}

http://www.aupress.ca/books/120146/ebook/99Z_Anderson_2008-Theory_and_Practice_of_Online_Learning.pdf

[9] Patricia Margaret Alexander, Towards reconstructing meaning when text is communicated electronically. PhD Thesis (Information Technology), University of Pretoria, Pretoria, South Africa, Aug 23, 2003, 392 pages.

<http://hdl.handle.net/2263/27367>

Full Copy

<https://repository.up.ac.za/bitstream/handle/2263/27367/Complete.pdf?sequence=13&isAllowed=y>

[10] Gary Woodill, "Rapid doubling of knowledge drives change in how we learn," Float, Jan 23, 2018.

<https://gowithfloat.com/2018/01/rapid-doubling-knowledge-drives-change-learn/>

[11] Alvin Toffler, Future Shock. New York, NY: Bantam Books, 1970 (re-issue 1984), 576 pages. {ISBN-13: 978-0553277371}

[12] Christopher M. Bishop, Pattern Recognition and Machine Learning. Cambridge, UK: Springer Science, 2004 (2nd ed.), 738 pages.

[13] Yingxu Wang, Du Zhang, and Witold Kinsner (eds.), Advances in Cognitive Informatics and Cognitive Computing. New York, NY: Springer, 2010, 316 pages. {ISBN: 978-3642160820}

[14] Yingxu Wang, "On cognitive informatics," in Proc. 1st IEEE Intern. Conf. Cognitive Informatics (Calgary, AB; 19-20 August 2002), pp. 34-42, 2002.

[15] Ning Zhong, Jiming Liu, and Yiyu Yao (eds.), Web Intelligence. New York, NY: Springer, 2010, 464 pages. {ISBN: 978-3642079368}

[16] Simon Haykin, José C. Principe, Terrence J. Sejnowski, and John McWhirter (eds.), New Directions in Statistical Signal Processing. Cambridge, MA: MIT Press, 2007, 514 pages.

[17] Simon Haykin, José C. Principe, Terrence J. Sejnowski, and John McWhirter (eds.), New Directions in Statistical Signal Processing. Cambridge, MA: MIT Press, 2007, 514 pages.

[18] Simon Haykin and Bart Kosko (eds.), Intelligent Signal Processing. Piscataway, NJ: IEEE Press, 2001, 573 pages.

[19] Richard G. Baraniuk, Emmanuel Candes, Robert Novak, and Martin Vetterli, "Compressive sampling" IEEE Signal Processing, vol. 25, no. 2, pp. 12-20, March 2008.

[20] Witold Pedrycz, Granular Computing: An Emerging Paradigm. New York, NY: Physica-Verlag HD, 2010 (2nd ed.), 403 pages. {ISBN: 978-3790824872}

- [21] Gregory W. Wornell, Signal Processing with Fractals: A Wavelet-Based Approach. Upper Saddle River, NJ: Prentice-Hall, 1996, 177 pages.
- [22] Holger Kantz, and Thomas Schreiber, Nonlinear Time Series Analysis. Cambridge, UK: Cambridge Univ. Press, 2004 (2nd ed.), 369 pages.
- [23] Julien Clinton Sprott, Chaos and Time-Series Analysis. Oxford, UK: Oxford Univ. Press, 2003, 507 pages.
- [24] Steven H. Strogatz, Nonlinear Dynamics and Chaos. Cambridge, MA: Westview/Perseus, 2000, 498 pages.
- [25] Robert L. Devaney, An Introduction to Chaotic Dynamical Systems. Menlo Park, CA: The Benjamin-Cummings Publishing, 1986, 320 pages.

Biography

Witold Kinsner is Professor in the Department of Electrical and Computer Engineering, University of Manitoba (UofM), Winnipeg, Canada, and Director of the Cognitive Systems Group. He was a co-founder of the first Microelectronics Centre in Canada, and was its Director of Research from 1979 to 1987. He is a Co-founding Member of the International Institute of Cognitive Informatics and Cognitive Computing (ICIC), Calgary, and the Canadian Engineering Education Association (CEEA). Since 1971, he has been very active at all the IEEE levels, including IEEE International, Region 7 (IEEE Canada), Society, Council, Section, Chapter, and Student Branch. He has organized many conferences, and has been on numerous editorial boards of journals and magazines. He is a member of 10 IEEE Societies and many committees. In 2013, Dr Kinsner was elected IEEE Canada President Elect 2014-15 and IEEE R7 Director/Delegate Elect 2014-15. He was IEEE Canada President 2016-17 and IEEE R7 Director/Delegate 2016-17, and he is now its Past President 2018-2019. In 2017, he was elected IEEE Vice President of Educational Activities Board for 2018.



He is part of the Symbiotic Autonomous Systems Initiative, FDC –IEEE, and co-author of the Initiative White Papers.

Email: witold.kinsner@umanitoba.ca

Are you ready for your “Digital Twin”?

Derrick de Kerckhove

Sommario

L'evoluzione verso sistemi sempre più autonomi che diventano sempre più integrati nella ns società e nella nostra vita, simbiotici, cambierà radicalmente il nostro modo di essere, aprendo nuove opportunità e ponendo nuove sfide. Un cambiamento che per impatto e dimensione potrebbe essere paragonato a quello che si è avuto nel Rinascimento. I Digital Twin in un certo senso sono sia una tecnologia abilitante sia un'icona di questo cambiamento.

L'articolo ne esplora le implicazioni e termina con un insieme di domande a cui forse non si riuscirà a dare risposte condivise ma che certamente stimoleranno la riflessione.

Among daily news about innovations in our non-stop collective creativity, two ground-breaking engineering concepts herald the new digital transformation. The first is called Symbiotic Autonomous Systems (SAS), a general category as well as a new evolutionary drive that brings together rapid advances in AI and Machine Learning applied to the worldwide accumulation of Big Data on everything. It will be the result of the concurrent evolution of machine augmentation and human augmentation converging on a symbiosis, that is, an intimate interpenetration between humans and machines or programs, a physical and mental co-existence in cyberspace. These engineering innovations will affect our bodies, our minds and the whole environment. Intelligent prosthetics, sophisticated robotics, decision-making algorithms, affective and sensitive digital assistants, and intelligent environments, well beyond today's idea of smart city, all this points to the formation not only of an entirely new way to occupy space and time, but also new experiences of one's perception and management of one's own selfhood.

Indeed, included among the considerations about SAS, the other life-changing innovation is the concept of “Digital Twin”. It was conceived a few decades ago as an engineering instrument that would augment machines with a digital double to monitor complex and expensive motors, turbines, and also robots. A machine's Digital Twin includes not only real-time reporting as it functions, but also the history of its maintenance, perhaps the occasional break-downs and,

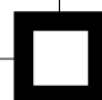
0

1

0

1

0



more importantly the source and coordinates of all parts suppliers and records of delivery and efficiency. In many cases, the digital twin allows to automate repairs just as it regulates normal functions. Today, however, the twinning concept is of interest not only to engineers, but also to town planners, military strategists, security systems operators, educators, trainers and a host of other complex operations managers. Planning is on course to extend and apply the concept to businesses, cities, regions and, why not, to whole countries. Digital Twin has made it to the top of Gartner's hype cycle last year and is not coming down quite yet because it is also destined to be applied to you and me. How come?

The idea of twinning you and me on line crowns thirty years of sketchy development that started with Apple's purely conceptual “Knowledge Navigator” (1987) and matured today on the market as Apple's Siri, Microsoft's Cortana, Samsung's Bixby, Google's Assistant, and particularly Amazon's Alexa, so far the most sophisticated one, in part because it integrates machine learning to learn about you as it manages, responds, sorts and stores your queries, and also because it is capable of expressing and responding to emotions. All these are great tools for daily use and the market's ambition is to make them your companions for life. The Digital Twin, however, is much more ambitious. It plans to represent you in every way and in all times, past, present and even future as it will be able to predict or propose moves on the basis of what it (he? she?) knows about you and what access it has to the world's databases and to top-notch analytics.

The truth is, we already have digital twins because we keep accumulating tons of data about ourselves, some deliberately in Facebook, Twitter and LinkedIn, others not quite so consciously in every move we take. Even mind-reading books are in store. It is entirely possible that eventually, as we read our novels on Kindle or Kobo, even our thoughts and our feelings will become data that will be grist to the mill for our digital twin:

“If Kindle was to be upgraded with face recognition software and biometric sensors, it would know how each sentence influenced your heart rate and blood pressure. It would know what made you laugh, what made you sad, what made you angry. Soon, books will read you while you are reading them.” (Yuval Noah Harari)©

Fresh off the press, so to speak, there is even news about simple headsets that can read your mind and either translate the sentences you form in your mind into speech, or send it to databases, ready to be analysed¹.

The proliferation of surveillance cameras in the public space of several countries, notably in China is adding accurate data about our whereabouts and our behavior. A recent study by Cornell University concretely demonstrates how machine learning can attack our privacy. The researchers applied basic algorithms (that is, less complex than those commercially used by, for example,

¹ (REF: January 29, 2019 from Science Daily: Engineers translate brain signals directly into speech: <https://www.sciencedaily.com/releases/2019/01/190129081919.htm>).

Facebook and Google), to identify people with blurred, pixelated, and encrypted images. They were able to "show how to train artificial neural networks to identify faces, recognize objects and handwritten figures" with an accuracy of 71%. On the other hand, human accuracy was 0.19%. At this point, we can't win. So what's at stake is a change of civilization much more radical than anyone has predicted and for which we are evidently not prepared. A very real possibility is that even as we are emptying the contents of our private self in Big Data, something like a digital twin will take its place to help us negotiate daily life in the digital realm.

The problem for the moment, however, is that we do not have one but myriad little twins scattered in the world's databases. Most of these micro profiles of ourselves tailored by government and institutions such as PRISM or Google are fragmented by the uses they are meant for. The government doesn't need to know your diet, but only if you represent a threat to the community or if you are not paying your taxes. Google wants to know what you read and builds your profile based on your search pattern and also on the kind of company you keep since it insists that your Gmail activities be coordinated with the other services the mammoth search engine gives you. In fact, since we generally do not know how, or how much or in what way these micro-profiles are constructed and used, I have grouped the lot of them under the general banner of the "digital unconscious", everything that is known about that you don't know and that has more effect on your choices and motivations than whatever Freud told you about mom and dad.

The point, of course, of having a DT, is to gather all this data together in an orderly way in such manner as to, first, give you a measure of control over it, and second, to endow you with a sort of "digital you" that, having access to everything digital, on and off line, knows a lot more than you do and especially more than you know or remember about yourself.

So what could your DT do for you? Markus Steer, advisor with SAP, writing for Digitalist Magazine puts the emphasis on health and medical care:

"Already today, doctors and insurance companies not only collect information about past & present treatments & checkups (including diseases of direct family members), but also actual habits, as smoking, regular work-out or extreme sports. Technology makes it possible to combine the different sources of information and elaborate a transparent Digital Twin of the individual."

The representation of an individual person, including personal data like weight, health data, activity tracking data, and medical treatment data can help to establish predictive alerts and guide people to healthier lifestyles. It will give better insights and transparency on the individual's health situation by offering more data points. Data shared with doctors will help define better preventive strategies and recovery plans and shrink health care costs.

Adds Steer: “In the business world, employee training plays an important role in time to market. Workers can be trained in the digital world first, leading to tremendous resource optimization for on-site training.”^[1]

There is a lot more to speculate about digital twins, as it will become everybody’s interface to navigate and negotiate for each one of us our relationships with the Internet and the whole digital universe, having access to current and integrated information about anything and everything. But there is one urgent question about our twin that brings us back to symbiotic autonomous systems (SAS). Indeed, how autonomous will it be? A major issue about the rapid development of ever more sophisticated technologies is that they are becoming autonomous. In fact, the more autonomy we yield to our machines, the less we can keep for ourselves. For a system to be autonomous, it has to:

- Have an independent agenda (goal, plans, policy and context)
- Demonstrate a capability to make independent decisions through deliberation
- Perform independent execution of actions

Says J.P. Morgan top AI expert, Manuela Veloso:

Systems already communicate wirelessly, drawing on data in the cloud, or [are] helped by remote teams. You can think of AI systems in constant symbiosis with everything else, with other information on the web, with other AI systems, with humans next to them, with remote humans. It becomes not a problem of developing self-contained AI systems, but an AI system that can recognize when it does not know, or when it needs more information, or when it thinks something with some probability but it’s not sure. It’s not that it can solve all the problems up front, but it can rely on all these other sources around.

The symbiosis comes from the fact that the boundaries between our auxiliary systems and our bodies and mind are continuously thinning down. Not only is technical evolution conceding autonomy to airplanes and automobiles, but to conversational robots and common digital assistants such as Alexa or Siri and soon intelligent environments and decision-making systems in medicine, law, finances and personal affairs.

The boundaries are fading largely because as people tend to delegate their basic cognitive faculties to such efficient services, our memory goes into our smartphone, our judgment and reasoning abilities are trusted to A.I. and even our imagination and creativity to automated designing, drawing, painting, writing (for poetry and standard letters, news, etc.). Even as we outsource our mental operations and allow our memory to be distributed in different databases, we are emptying ourselves of the contents and cognitive strategies that centuries of patient accumulation and refinement had included within the privacy of our psyche.

Our IEEE research team identified many legal and security, as well as ethical issues emerging from the rapid development of new generation expert systems that are already replacing human skills in labor and professional competence in law, medicine, business and arbitration.

For example, IBM has already introduced Watson into the legal field, to cope with the complex structure of legal knowledge contained in the federal statutes, regulations, treaties, contracts and jurisprudence. Some algorithms are already being used to determine prison sentences. If we start trusting algorithms to make decisions, who will have the final word on important decisions? Will it be humans, or algorithms? One problem with algorithms is that, different from both expert systems and standard neural networks that can be analysed for how they process data, they begin being clearly traceable, but very quickly they escape the control of the programmers. AI has "a mind of its own", not necessarily a mal-intentioned one, but, at some point, depending on how much autonomy is given to it, a mind that can get off the track, sooner or later. Manuela Veloso is fully aware of this problem:

"So we are working on the ability for these AI systems to explain themselves, while they learn, while they improve, in order to provide explanations with different levels of detail. We want to interact with these robots in ways that make us humans eventually trust AI systems more. You would like to be able to say, "Why are you saying that?" or "Why are you recommending this?" Providing that explanation is a lot of the research that I am doing now, and I believe robots being able to do that will lead to better understanding and trust in these AI systems."^[2]

Another problem is that algorithms being extremely efficient and having access to infinitely more data, they beat humans at a growing number of life/death, or less dramatic estimates, diagnoses, legal rights and wrongs, financial advising etc. They also replace brain and brawn for not so menial jobs. It is said that by 2030, 45 % of the present workforce in advanced economies is going to be taken by algorithms and robots.

So beside jobs, what are the potential victims of algorithms?

Autonomy: the more decisions are made by machines, the less freedom of decision and movement is left for people;

Individual property and control of thought: we are no longer exclusive owners of our thinking because I.A. is perfectly capable of retracing it, even in real time;

Psychological interiority: I'm emptying myself, first because instead of building my inner self, I pour my self in social media, then because the machine is acquiring expanding and extending my internal cognitive functions, including memory, intelligence and imagination;

Reputation: the era of ubiquitous public and private surveillance and one-way transparency puts us naked in the street.

And so on. The real question is how do humans behave under such condition of vulnerability.

Not only does the adoption of aggressive new technologies always raise ethical questions, some of them such as the appearance of the printing press, or now the digitization of human culture, demand a radically new ethical order. The Renaissance witnessed a brutal redefinition of what was meant to be human

during the painful transition between a predominantly oral and communal religious authority to an individualistic humanist social and political order. The spread of alphabetic literacy individualized people by providing them with a private – and relatively easy – access to language. By learning to read silently, they could privatize not only their control over language, but also interiorize their conscious activities and become private persons. This fact changed their relationship to themselves, to their beliefs and to other people, family, friends and foes. While transiting from the so-called culture of “shame” to that of “guilt”, the object of their personal responsibility shifted from “the other” to the “self”. This is also the origin of puritanism, that is the deep sense of one’s responsibility to one’s destiny in this world and in whatever other realm they believed in.

And this, fundamentally is what is changing, not to say reversing today. As people are ever more exposed to continuous monitoring by automated electronic systems, and while, in some countries, behavior itself is controlled by algorithms, responsibility shifts away from the self to the now almost self-organizing whole social order, including - and perhaps eventually prioritizing – the care of the environment.

Shouldn’t symbiotic autonomous systems always be developed with their after-effects on people and the environment in mind? Shouldn’t large scale predictive analytics be applied to each innovation before implementation? Even as algorithms and A.I. take the lead in introducing a generalized symbiosis between individuals and the environment itself, shouldn’t an ethical dimension be consciously included in their programming?

Among the most urgent questions, is how SAS need to take account of the epistemological change that is going on right now because of technology. People are being emptied of their psychological content but hardly noticing it because they labor under the illusion that, apart from additions to our technical capabilities, human nature, their own in particular, is the same as it always was. “Human nature” changed before and it is changing again, so much so that the unrecognized transition from humans developing from within to projecting (or rather letting go of) one’s identity on line externally could bring one to doubt that there is such thing as human nature at all. For example, is it not conceivable that the evolution of our “digital twins” will eventually become the next modality whereby people negotiate their relationships with SAS and the total environment, as opposed to what westerners at least still consider the activity and the exclusive property of internal memory, intelligence and judgment? Assuming symbiosis is completed within the next thirty years, will humans still benefit from any capacity to resist intellectually, let alone politically or even emotionally?

Although it is suggested in our latest White Paper that one among the ways to offset the complete take-over of human decision-making by A.I. and algorithms, would be to genetically modify the human genome to increase its intelligence on a par with Artificial General Intelligence (AGI), the predictable future is that humans will become increasingly vulnerable to SAS’ domination without allowing enough time to evolve that sort of recourse. We cannot stop the exponential growth and sophistication of our machines, but we can decide to infuse them *algorithmethically* with genuine human survival and environmental priorities

and monitor and control worldwide how they are applied. An ethical taskforce of SAS projections would want to address the particulars of that issue and that will be a large part of our next issue.

For example, regarding the potential social injustice that would occur in giving access to human augmentation by technological or biological means, one could invoke regulation obtaining in sport competition, but it isn't quite the same thing to prevent doping in a cycling event and to impede school and career privileges to enhanced humans. How does one regulate for social justice in income and means disparities in a livable SAS environment? And, at the same time, can we allow regulation to stand in the way of the full realization of enhanced human potential. Observing the rapid proliferation of Brain Computer Interface (BCI) and slower but continuous advance of Computer Brain Interface (CBI) technologies, it appears that connecting brains to the internet and eventually accessing directly and pertinently its enormous contents is almost a foregone conclusion, if not seamlessly at first, perhaps via tweaking our digital twin. Wouldn't then be a need to prevent intentional or accidental harm through thoughtless activity by re-programming the Internet itself?

Finally, SAS predicts the gradual evolving of super-organisms consisting of the task-oriented association of indefinite numbers of components, human and algorithmic. This prediction is congruent with observations above regarding how humans may become more and more defenseless about how SAS takes over all important decision-making. Should such a possibility be left to teleological self-organization or should it become the object of an international political decision-making?

Perhaps the biggest problem with algorithms is that they are about to keep all of us under strict control. I have called that datacracy. What I mean by it is that data surveillance that is presently practiced, for example, in Singapore, can quickly turn into total control by data with political and institutional leverage. When I say 'practiced' I mean, openly done, with public assent, just as the social credits project in China, effectuate the same, if not higher level of control, in full knowledge of the population, and, according to surveys and interviews reports, not even grudgingly.

This is the difference between the west and the east here. In the west, business and government do more or less the same things, but discreetly and more on the side of business than that of government (except, of course in neo-totalitarian countries). And all of it pretty well under cover. You would think, after so many revelations by so many whistleblowers, that people should probably begin to get the idea. In matter of fact, they probably do but don't really care; they realize it cannot be stopped. It's destiny. In the east, on the other hand, community spirit and a desire to share a safe and prosperous environment with everyone overrides the priorities of the individual.

Ultimately, algorithms might get rid of government altogether. In a full-fledged Symbiotic Autonomous Environment, all parameters are taken under consideration, even those that measure and locate where and how we treat animals, or how much recycling we may or may not have done. The new ethics, if however, we eventually develop it, will have to be grounded in protecting the

environment. In a mature datacracy, we do not turn into servo-mechanisms of 'The Machine', we do not serve an oligarchy of power-hungry technological monarchs, we do not rise against each other for whatever advantage we can gain. In a mature datacracy, humans leave machines do their job, which is to take care of our problems, all of them at once and all the time. That is what an environmentally aware general AI can eventually accomplish. But we are still a fair bit away from that...

Since this paper is destined to close the present issue of Mondo Digitale, I want to conclude with an invitation to reflect on some basic questions we might need to ask ourselves even as we begin a new industry of digital twins.

What format could/should industry devise for our digital twin (augmented Digital Assistant applications – different 'skins' – long term – occasional/permanent)?

Could our DT become a fashion item or an art form?

How many DTs should a single person be allowed (secure identifiers – formal ownership contractual policy)?

What content should DTs integrate (lifelogging – punctual – real-time/continuous updating – open or restricted database access)?

What kind of connectivity would be most useful (to owner - to other twins)?

What software capability should be programmed into our DT (various levels Data Analytics – descriptive – predictive - prescriptive)?

What kind of regulation – if any – should DTs be subject to (responsibility – accountability –

What level of autonomy should the DT be granted (unsupervised – supervised – accountable)?

What order of ethical priorities should be mandatory for developing and using DTs (ethical behavior – manners – being answerable)

What legal status should DTs be granted?

What should be done with one's DT after one's demise?

What roadmap should be considered for developing DTs?

Biography

Derrick de Kerckhove, sociologo, giornalista e direttore scientifico di Media Duemila. Ha diretto dal 1983 al 2008 il McLuhan Program in Culture & Technology dell'Università di Toronto.

È autore di "La pelle della cultura e dell'intelligenza connessa" ("The Skin of Culture and Connected Intelligence") e Professore Universitario nel Dipartimento di lingua francese all'Università di Toronto. Già docente presso il Dipartimento di Scienze Sociali dell'Università degli Studi di Napoli Federico II dove è stato titolare degli insegnamenti di "Sociologia della cultura digitale" e di



Are you ready for your "Digital Twin"?

"Marketing e nuovi media". È supervisor di ricerca presso il PhD Planetary Collegium T-Node. È membro dello Steering Board della Symbiotic Autonomous Systems Initiative e co autore dei due White Paper prodotti.

Email: dekerckh@gmail.com