

Astrazione automatica di regole per sistemi IoT: un approccio rivolto ad utenti non esperti

Studente: Ballati Fabio

Relatori: Fulvio Corno, Luigi De Russis

Introduzione

Internet è ormai da diverso tempo uscito dai confini degli schermi dei computer e anche da quelli dei tablet e degli smartphone, e si è diffuso come un liquido in tutto ciò che ci circonda.

Internet of Things (IoT) è il neologismo riferito all'estensione di internet al mondo degli oggetti e dei luoghi concreti. Attraverso chip e sensori inseriti al loro interno, gli oggetti sono in grado di interagire tra loro e con la realtà circostante. Così il mondo fisico può essere quasi interamente digitalizzato, monitorato ed in molti casi virtualizzato. Grazie ai dispositivi IoT la nostra casa si chiuderà da sola quando ci allontaniamo, saremo guidati dalla nostra auto al distributore di benzina (aperto) più vicino, potremo tenere sotto controllo la nostra salute attraverso gli smartwatch allacciati al polso.

Programmazione basata su regole

Lo sviluppo dei sistemi IoT implica l'incremento del numero di dispositivi che interagiscono con ciascun utente, ogni giorno ed in ogni luogo. Un crescente bisogno di personalizzare il comportamento di questi dispositivi da parte di utenti non esperti è ora gestito attraverso servizi che fanno uso della programmazione basata su regole, come IFTTT (if-this-then-that).

Un esempio di regole create attraverso IFTTT è:

IF "hai pubblicato un nuovo post" su "Facebook"
THEN "aggiungi il testo al documento" presente in "Google Drive"

IF "hai pubblicato un nuovo post" su "Facebook"
THEN "aggiungi il testo al file di testuale" presente in "Dropbox"

Nella prima regola (anche in Figura 1), l'evento chiamato "trigger" è "hai pubblicato un nuovo post", il "canale di trigger" su cui è esso viene eseguito è "Facebook". L'evento, che viene eseguito allo scattare del trigger, chiamato "azione", è "aggiungi il testo al documento". Il "canale di azione" su cui essa viene eseguita è "Dropbox".

Obiettivo della tesi

Studi recenti hanno mostrato come piattaforme simili a IFTTT forzino gli utenti a lavorare a livello dello specifico dispositivo ("basso livello"), mentre l'utente preferirebbe fare uso di un livello leggermente più astratto, nel quale le regole sono più generali e comprensibili ("alto livello").

L'obiettivo della tesi è quello di **definire, progettare e validare** una metodologia per il mapping automatico delle regole di basso livello (come quelle di IFTTT) su "regole di alto livello", più astratte.

Il processo di **definizione e progettazione**, prevede la creazione di canali di alto livello in cui saranno suddivisi trigger ed azioni di basso livello. A partire da questi canali un algoritmo esegue il mapping tra le regole di basso livello, prelevate dal dataset contenente le regole pubbliche di IFTTT illustrato in seguito, e le regole di alto livello.

Il processo di **validazione** esplorerà le seguenti domande: quante regole di basso livello esistenti possono essere tradotte con successo in regole di alto livello e quante no? Quali sono le regole di associazione che determinano la somiglianza (stesso canale di alto livello) tra le regole di basso livello? Qual è la percezione dell'utente relativa alle regole astratte tradotte, quando le si confronta con le corrispondenti regole di basso livello?

Le regole di alto livello

Le regole di alto livello sono strutturate come le regole di basso livello.

La differenza è data dalla possibilità di poter scegliere contemporaneamente in una stessa regola, per uno stesso evento, più servizi attraverso i quali esso eseguire l'evento desiderato.

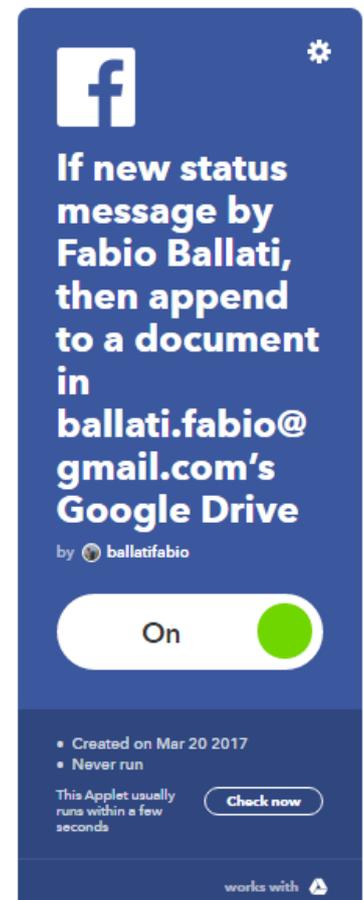


Figura 1

Un esempio di regola di alto livello è:

IF "hai pubblicato un nuovo contenuto testuale" su uno o più tuoi servizi SOCIAL (Facebook)
THEN "aggiungi il testo del contenuto al documento" su uno o più tuoi servizi CLOUD (Google Drive, Dropbox)

Nella regola presentata (anche in Figura 2), il "trigger di alto livello" è "hai pubblicato un nuovo contenuto testuale", il "canale trigger di alto livello" su cui esso viene eseguito è "SOCIAL". Per le regole di alto livello è introdotto il concetto di "servizio". Nell'esempio, il servizio relativo al trigger è "Facebook". L'"azione di alto livello" che viene eseguita allo scattare del trigger è "aggiungi il testo del contenuto al documento". Il "canale azione di alto livello" su cui essa viene eseguita è "CLOUD". Il servizio per tale azione è dato dalla coppia "Google Drive", "Dropbox".

Le regole di alto livello permettono di selezionare più servizi (Google Drive e Dropbox nell'esempio) per ciascun evento. Il risultato è quello di collegare più dispositivi/ servizi aventi funzionalità simili utilizzando una sola regola. Si ottiene un risparmio relativo al numero di regole create; esse non sono più due come nel paradigma di basso livello, bensì una.



Il Dataset

È stato analizzato il dataset (del progetto UPOD ("Universal Programming Of Devices", ricerca accademica della Brown University) contenente le regole pubbliche create dagli utenti di IFTTT. Il dataset comprende 303 azioni (presenti in 143 canali azione) e 697 diversi trigger (divisi in 177 canali trigger). Le regole totali sono 224,590. Il dataset è relativo al settembre 2015.

Definizione e Progettazione della metodologia di mapping automatico basso/ alto livello

La definizione della metodologia di mapping è basata sullo studio del dataset. Sono state individuate correlazioni e similitudini tra gli eventi in esso contenuti, i quali hanno portato alla definizione del concetto di "canale di alto livello". Questa è un'etichetta, pensata a partire dal linguaggio quotidiano delle persone, alla quale sono associati più *eventi di alto livello* appartenenti ad un ambito simile (quello del canale). Sono stati individuati 13 canali di alto livello per i trigger e 7 per le azioni.

- Esempi di canali di alto livello presenti sia per i trigger che per le azioni sono:
 - Social**, relativo i contenuti condivisi con altre persone sulla rete; **Cloud**, relativo i contenuti memorizzati in rete; **Smart-Home**, relativo i dispositivi collegati ad internet presenti in casa; **Dispositivi personali** ...
- Esempi di canali di alto livello presenti solo per i trigger sono:
 - Posizione**, relativo alla posizione geografica attuale; **Tempo Atmosferico**, **Data&Tempo**, **Interessi** ...

All'interno di ciascun canale, a partire dal suo ambito, sono definiti più "eventi di alto livello" (Figura 3), pensati a partire dal linguaggio quotidiano delle persone. A questi sono stati associati infine i "servizi", cioè le applicazioni pratiche sulle quali l'evento verrà eseguito.

I servizi permettono di ottenere il parallelismo (la concreta possibilità di mapping) con le regole di basso livello.

Da ultimo si è constatato come una regola di alto livello eseguita tramite più servizi, fosse equivalente, da sola, a più regole di basso livello. La fase di progettazione prevede di implementare un algoritmo avente tale funzionalità: dato un insieme di regole di basso livello, creare (un minore numero di) regole di alto livello mantenendo inalterate le funzionalità iniziali.

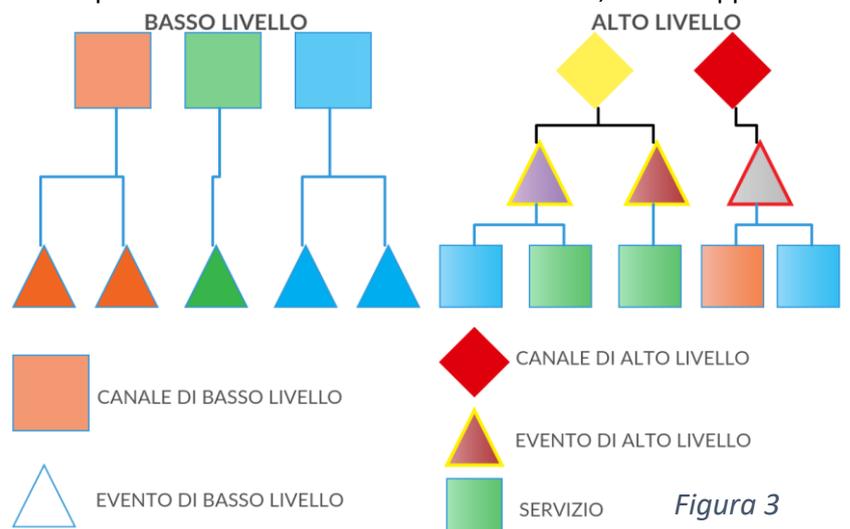


Figura 3

La **progettazione** della metodologia di mapping ha visto la realizzazione di un algoritmo che, per ciascun utente separatamente, esegue il mapping delle regole di basso livello da esso create, in regole di alto livello e tiene memoria di quante regole di alto livello siano state create per coprire l'insieme di regole di basso livello realizzate dall'utente. L'algoritmo utilizza la seguente logica:

- Si identificano insiemi di regole i cui trigger appartengono allo stesso canale di alto livello di trigger, lo stesso deve valere per le azioni.
- In questi gruppi si individua il gruppo di regole più numeroso possibile, che siano mappabili in un'unica regola di alto livello. La ricerca del gruppo di regole è fatta usando una struttura dati che è a conoscenza per ciascun trigger di quali e quante siano le azioni di un certo canale, e viceversa per le azioni.
- L'algoritmo elimina le regole di basso livello, già mappate in una di alto livello, e ripete l'operazione di ricerca del gruppo di regole con le caratteristiche sopra descritte

Come output l'algoritmo fornisce, per ciascun autore, le regole di basso livello iniziali, le regole di alto livello create, la differenza della numerosità tra alto e basso livello.

Validazione della metodologia di mapping automatico basso/ alto livello

Il processo di validazione della metodologia di mapping risponderà alle domande sopra elencate soffermandosi su tre aspetti: il risparmio in termini di numero di regole create attraverso il paradigma di alto livello (valutato usando l'algoritmo descritto - i), il miglioramento della user experience e il miglioramento della usabilità nell'utilizzo di una web application facente uso del paradigma di alto livello (valutato attraverso un test utente - ii).

i - Applicando l'algoritmo sopra descritto sul dataset a disposizione, emergono i seguenti dati:

- Per autori che hanno creato più di due regole, sono state mappate le 115872 regole di basso livello su 97383 regole di alto livello, con un risparmio medio per autore del 15,96% sul numero di regole realizzate.
- Per autori che hanno creato più di dieci regole, sono state mappate le 55681 regole di basso livello su 41774 regole di alto livello, con un risparmio medio per autore del 24,98% sul numero di regole realizzate.
- Per autori che hanno creato più di due regole e ne hanno risparmiata almeno una attraverso il paradigma di alto livello, sono state mappate le 42338 regole di basso livello su 28366 regole di alto livello, con un risparmio medio per autore del 33% sul numero di regole realizzate.

La selettività rispetto agli autori ha lo scopo di analizzare coloro che fanno un utilizzo almeno modesto di IFTTT. I dati numerici dimostrano come un paradigma di regole di alto livello sia vantaggioso per l'utilizzatore.

ii - Il test utente si compone di tre scenari, ciascuno con una serie di task. Gli otto utenti che hanno partecipato al test hanno svolto gli scenari con entrambi i paradigmi, utilizzando una apposita applicazione web. Al fine di valutare usabilità ed user experience, dopo ciascuno scenario i tester hanno dato una risposta numerica (1 non facile – 5 decisamente facile), per entrambi i paradigmi, alla domanda "È stato facile svolgere lo scenario proposto?".

L'andamento delle risposte è stato nettamente a favore del paradigma di alto livello (Figura 3), anche se non nel 100% dei casi. A questo è stata attribuita una valutazione di 5 relativamente a dodici scenari su ventiquattro totali, mentre per il basso livello solamente quattro scenari hanno ricevuto valutazione 5. Similmente, per quanto riguarda il paradigma di alto livello, zero scenari hanno ricevuto 1 come valutazione ed un solo scenario ha ricevuto 2. Sono in tutto dieci le risposte indicanti i valori 1 e 2 nel caso del paradigma di basso livello. Inoltre utilizzando il paradigma di basso livello, non tutti gli scenari sono stati completati nel modo corretto.

I commenti orali rilasciati dalla maggioranza dei tester, a proposito di come sia stato maggiormente agevole svolgere gli scenari con il paradigma di alto livello, confermano quanto emerso dai dati: usabilità ed user experience migliorano con un paradigma che fa uso di un livello di astrazione leggermente più alto.

Il test ha dato un'ulteriore dimostrazione relativa il risparmio in termini di regole create. I tester che hanno eseguito nel modo corretto lo scenario, hanno utilizzato 20 regole di basso livello oppure 9 di alto livello.

Conclusione

Il paradigma di alto livello, i cui canali ed eventi sono ideati a partire dal modo di pensare e dal linguaggio delle persone che ne usufruiscono, ha la forza di unire ai vantaggi puramente pratici dimostrati, una maggior valorizzazione e gradevolezza dell'esperienza utente, volta anche a rendere più semplice ed accessibile l'utilizzo del servizio con lo scopo di agevolare la gestione quotidiana dei dispositivi nell'attuale mondo IoT.