

UNIVERSITÀ DI PADOVA



FACOLTÀ DI INGEGNERIA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
Corso di Laurea Magistrale in Ingegneria delle Telecomunicazioni

Tesi di Laurea

**POLITICHE DI SCHEDULING OTTIME
PER SENSORI RADIO ALIMENTATI
DA FONTI DI ENERGIA RINNOVABILI**

Relatore: Prof. Michele Rossi

Laureando: Marco Fasolo

Anno accademico 2010/2011

Padova, 11 Luglio 2011

*A mamma, papà,
Elena e Roby.*

Indice

Indice	i
Sommario	1
1 Introduzione	3
1.1 Struttura della tesi	7
2 Processi decisionali di Markov	9
2.1 Introduzione ai processi decisionali di Markov	9
2.2 Definizioni e notazione	10
2.2.1 Istanti di decisione	10
2.2.2 Stati	11
2.2.3 Azioni	11
2.2.4 Probabilità di transizione	11
2.2.5 Costi	12
2.2.6 Processo decisionale di Markov	12
2.2.7 Regola di decisione	13
2.2.8 Politica	13
2.3 Criterio di ottimizzazione	14
2.3.1 Assunzioni	14
2.3.2 Costo medio per slot	15
2.3.3 Costo totale α -scontato	16
2.4 Ottimizzazione secondo il criterio del costo medio per slot	17
2.4.1 Definizioni utili sulle catene di Markov a stati finiti	18

2.4.2	Relazione tra costo medio per slot e costo totale α -scontato . . .	19
2.4.3	Politica Blackwell ottima	26
2.4.4	Equazioni di ottimalità	35
3	Modello del sensore radio	39
3.1	L'idea	39
3.2	Assunzioni	42
3.3	Obiettivi	43
3.4	Batteria	44
3.4.1	Livello di guardia	45
3.5	Fonte di energia rinnovabile	46
3.5.1	Modello generale	46
3.5.2	Modello generale a stato singolo	48
3.5.3	Energia solare: modello a due stati	49
3.5.4	Energia solare: modello a più stati	51
3.6	Radio, sensore, unità di controllo e buffer	51
3.6.1	Radio e sensore indipendenti	54
3.6.2	Radio e sensore mutuamente esclusivi	55
3.6.3	Buffer	55
3.6.4	Modalità di funzionamento	56
3.7	Modello complessivo per il sensore radio	57
3.7.1	Istanti di decisione	57
3.7.2	Stati	57
3.7.3	Azioni	58
3.7.4	Probabilità di transizione	61
3.7.5	Costi	63
3.8	Osservazioni conclusive sul modello	67
4	Algoritmo di ottimizzazione	69
4.1	Classificazione dei processi decisionali di Markov	69
4.1.1	Classi	70

4.1.2	Metodo di classificazione	72
4.2	Processi comunicanti in senso lato e costo medio per slot	73
4.2.1	Proprietà del modello proposto per il sensore radio	75
4.3	Procedura di ottimizzazione	76
5	Simulazioni, analisi e risultati	87
5.1	Parametri del modello	87
5.1.1	Consumi energetici	88
5.1.2	Batteria	89
5.1.3	Fonte di energia rinnovabile	90
5.1.4	Buffer, traffico esogeno e locale	93
5.2	Risultati ottenuti mediante simulazioni	94
5.2.1	Politica ottima e altre politiche	95
5.2.2	Utilizzo del parametro β	98
5.2.3	Consumo energetico dell'attività di sensing	101
5.2.4	Energia fornita dalla fonte rinnovabile	103
5.3	Risultati ottenuti mediante analisi delle politiche ottime	104
5.3.1	Modello a stato singolo per la fonte di energia rinnovabile	105
5.3.2	Livello di guardia della batteria	108
5.3.3	Modalità di funzionamento	109
5.3.4	Modello a due stati per la fonte di energia rinnovabile	111
5.3.5	Modello a due stati vs modello a stato singolo	112
5.3.6	Azioni specificate dalla politica ottima	117
5.4	Riepilogo dei risultati ottenuti	119
6	Conclusioni e sviluppi futuri	121
	Bibliografia	127

SOMMARIO

Nell'ultimo decennio le *reti di sensori wireless* o *WSN (Wireless Sensor Network)* hanno riscosso un notevole successo e sono oggi utilizzate in molteplici ambiti applicativi, soprattutto per il monitoraggio di particolari fenomeni fisici o ambientali. Le loro peculiarità consistono specialmente nel basso costo dei nodi che le costituiscono, chiamati *sensori radio*, e nella possibilità di dislocare quest'ultimi in aree pericolose o difficilmente raggiungibili dall'uomo.

I sensori radio di una WSN sono generalmente alimentati mediante batterie di capacità finita e il problema principale che si deve affrontare nella progettazione di una rete di questo tipo è l'autonomia limitata dei nodi che la costituiscono. Una soluzione, sulla quale si sono concentrate le attività di ricerca negli ultimi anni, è quella di sfruttare fonti di energia rinnovabili, come l'energia solare, per ricaricare la batteria di ciascun dispositivo. L'utilizzo di fonti di energia rinnovabili, tuttavia, pone delle importanti problematiche legate alla impossibilità di prevederne deterministicamente il comportamento: il contributo energetico fornito a ciascun dispositivo della rete è aleatorio nel tempo e variabile anche in funzione della posizione occupata.

L'oggetto di studio di questa tesi, che si colloca nel contesto appena delineato, è un generico sensore radio alimentato mediante una fonte di energia rinnovabile. Il principale contributo del lavoro svolto consiste in un modello probabilistico in grado di descrivere in modo formale e rigoroso un sistema di questo tipo. Nello specifico, il sensore radio viene rappresentato mediante un *processo decisionale di Markov*. Il modello proposto consente di formulare e trattare dal punto di vista matematico il problema della determinazione della *politica di scheduling ottima* che il sensore radio deve seguire per sfruttare nel modo più efficiente possibile la fonte di energia a sua disposizione. Ciò significa individuare quali siano le azioni ottime che di volta in volta il dispositivo deve compiere per raggiungere le massime prestazioni. Il pregio del modello realizzato è quello di essere generale e parametrizzabile: per rappresentare un particolare sensore radio alimentato da una specifica fonte rinnovabile in determinate condizioni operative è sufficiente fissare opportunamente i parametri del modello.

Oltre al modello per il sensore radio, si propone anche un *algoritmo di ottimizzazione* ad esso applicabile che consenta di determinare, una volta fissati i parametri, la corrispondente politica di scheduling ottima. Si giunge, così, alla realizzazione di un *ottimizzatore*, ottenuto implementando in C++ sia il modello che il relativo algoritmo di ottimizzazione.

Sfruttando l'ottimizzatore così costruito, viene anche verificata, mediante alcune simulazioni, la validità del modello e dell'algoritmo proposti. Si studiano poi alcune caratteristiche interessanti delle politiche ottime calcolate al variare dei parametri.

1

Introduzione

Lo sviluppo tecnologico recente nel campo delle comunicazioni wireless, dei sistemi microelettromeccanici e dell'elettronica digitale ha portato le cosiddette *reti di sensori wireless* o *WSN (Wireless Sensor Network)* ad assumere un ruolo di cruciale importanza sia nel mondo della ricerca che in ambito applicativo. Una WSN è sostanzialmente una rete distribuita di dispositivi radio che collaborano tra loro nel monitoraggio di un particolare fenomeno fisico o ambientale di interesse. Ciascun nodo della rete, comunemente chiamato *seniore radio*, è un dispositivo di dimensioni molto ridotte, con limitate capacità computazionali e di memorizzazione, a basso consumo di potenza e relativamente economico. Tipicamente è dotato di un sensore capace di rilevare e misurare determinate grandezze d'interesse nell'ambiente circostante, di una radio per la trasmissione e ricezione dei dati, di un'unità di controllo e di una fonte di alimentazione.

Le WSN, sebbene inizialmente siano state utilizzate soprattutto per scopi militari, trovano oggi numerose applicazioni in ambiti molto diversi; quelle più interessanti e diffuse prevedono l'utilizzo di dispositivi alimentati a batteria e quindi in grado di operare in modo autonomo. Ciò che rende queste reti particolarmente attraenti sono soprattutto le dimensioni ridotte e il costo limitato dei sensori radio nonché la possibilità di collocare i dispositivi in zone difficilmente raggiungibili o pericolose. Tipico è l'utilizzo delle WSN per il monitoraggio ambientale, ad esempio per rilevare l'attività sismica in prossimità di vulcani attivi [31] o per studiare l'habitat e i flussi migratori di animali selvatici [19], oppure per il controllo di infrastrutture quali ponti [17] e linee ferroviarie [6] o, ancora, per la localizzazione di veicoli e quindi per il controllo del traffico o della velocità [14].

Il problema principale che ci si trova ad affrontare quando si vuole realizzare una rete di sensori alimentati a batteria è la capacità finita di quest'ultima. Ciascun nodo della rete può operare solo per un tempo limitato ovvero fino a quando la carica a disposizione non si esaurisce completamente. Conseguentemente, le applicazioni che

utilizzano tali reti devono essere anch'esse di durata finita oppure sono necessari costi e complessità aggiuntivi per cambiare o ricaricare periodicamente le batterie. D'altro canto, per disporre di sensori radio ad alte prestazioni, è richiesto naturalmente un elevato dispendio di energia che implica un tempo di vita ridotto. Quindi si rivela necessario raggiungere sempre un compromesso tra prestazioni e durata della batteria.

Negli ultimi anni molte attività di ricerca sono state dedicate allo studio di possibili strategie per massimizzare il tempo di vita dei nodi alimentati a batteria di una WSN. Diverse sono state le soluzioni proposte: protocolli efficienti dal punto di vista energetico per l'accesso al mezzo [32], per il routing [9], per la disseminazione [10] e memorizzazione dei dati [7]; tecniche adattive per la rilevazione dei dati ambientali [18]; ridondanza nel posizionamento dei nodi [16]. Tutti questi espedienti, tuttavia, riescono solo a mitigare ma non a risolvere completamente il problema di fondo dato che l'energia a disposizione per ciascun dispositivo rimane comunque finita.

Una strada alternativa che si è iniziato a percorrere negli ultimi anni è quella di utilizzare fonti di energia rinnovabili per alimentare i nodi di una WSN. L'idea è quella di convertire l'energia raccolta dall'ambiente circostante in energia elettrica che possa essere impiegata per ricaricare la batteria di ciascun dispositivo. Le sorgenti di energia che possono essere sfruttate in questo senso sono molteplici: il sole, il vento, le onde radio, le vibrazioni e i movimenti umani sono solo alcuni esempi. Se l'energia accumulabile in questo modo è sufficientemente elevata e disponibile con continuità o con una certa periodicità, il sensore radio potenzialmente può operare per un tempo infinito. Questa strategia, quindi, sembra essere adatta per risolvere con successo il problema del tempo di vita finito dei nodi di una WSN, soprattutto nelle situazioni in cui la sostituzione della batteria o la ricarica manuale sono impraticabili ed è quindi richiesta una completa autonomia dei nodi. Per questo motivo recentemente sono stati progettati e realizzati diversi sensori radio capaci di sfruttare le fonti di energia presenti nell'ambiente circostante. In particolare è stato prediletto l'utilizzo dell'energia solare essendo facilmente disponibile e anche conveniente dal punto di vista dell'efficienza di conversione; alcuni esempi di dispositivi dotati di pannello solare sono: *Prometheus* [11], *HydroWatch* [28] e *Heliomote* [23]. Chiaramente sono stati

costruiti anche altri prototipi in grado di utilizzare diverse fonti di energia rispetto a quella solare; una panoramica completa è fornita in [26]. Ciò che differenzia le varie implementazioni è la tecnologia adottata per prelevare l'energia dall'ambiente e per immagazzinarla nel dispositivo.

L'utilizzo di fonti di energia rinnovabili per ricaricare la batteria dei nodi di una WSN introduce tuttavia le importanti problematiche legate alla natura imprevedibile e aleatoria di tali sorgenti di energia. Si pensi ad esempio all'energia solare: l'apporto energetico varia non solo in base al momento della giornata ma anche secondo le particolari condizioni ambientali che naturalmente non sono prevedibili deterministicamente. L'attività del dispositivo e quindi le sue prestazioni nel tempo dipendono strettamente dall'evoluzione della sorgente di energia.

Il lavoro contenuto in questa tesi si colloca nello scenario appena descritto. L'oggetto di studio è un sensore radio capace di acquisire energia dall'ambiente e di immagazzinarla localmente nella batteria. Il primo problema che viene affrontato nella tesi è quello di rappresentare questo tipo di sistema in modo formale mediante un modello matematico il più semplice e generale possibile. A tal fine si individuano le componenti fondamentali che caratterizzano il dispositivo: la *radio* per la ricezione e la trasmissione delle informazioni da e verso gli altri nodi, il *sensore* o *unità di sensing* per la rilevazione del fenomeno d'interesse, due *buffer* per la memorizzazione delle informazioni provenienti dalla radio e dal sensore, la *batteria* necessaria per l'alimentazione e una *sorgente di energia* che ricarica la batteria. Si considerano, poi, le due operazioni principali che un nodo di una WSN è chiamato ad eseguire: l'*inoltro* delle informazioni provenienti da altri nodi della rete e la *rilevazione* dei dati desiderati dall'ambiente circostante. L'attività del sensore radio è gestita dall'*unità di controllo* che stabilisce di volta in volta l'azione che deve essere compiuta. Il modello che si propone con questa tesi è un modello probabilistico che rientra nell'ambito generale della *programmazione dinamica stocastica* e rappresenta il dispositivo mediante un *processo decisionale di Markov* a tempo discreto [1, 2, 21, 25].

Il modello realizzato viene sfruttato per determinare quali siano, e in che modo possano essere raggiunte, le migliori prestazioni del sensore radio compatibilmente

con la batteria e la fonte di energia rinnovabile a disposizione. Con questo intento vengono utilizzati i potenti strumenti di ottimizzazione forniti dalla teoria dei processi decisionali di Markov. Al variare dei parametri in gioco, si individuano quali siano le azioni ottime che il dispositivo deve compiere durante la sua evoluzione al fine di massimizzare le prestazioni, ovvero il numero di informazioni trasmesse provenienti dall'attività di rilevazione e inoltro. Ciò che viene determinato sono le *politiche di scheduling ottime* che indicano in che modo debba agire il sensore radio in un dato istante in base alla condizione in cui si trova in quel momento. L'aspetto interessante è che l'ottimizzazione viene eseguita *tenendo conto* dell'evoluzione stocastica del livello di carica della batteria e della sorgente che la alimenta: il calcolo delle politiche ottime è dunque *consapevole* dell'apporto energetico aleatorio su cui il dispositivo può contare.

Il pregio del modello proposto è la sua generalità. Teoricamente può essere adottato per rappresentare un nodo alimentato da una qualsivoglia fonte di energia rinnovabile e dotato di un qualsiasi tipo di batteria; è sufficiente impostare in modo opportuno i parametri che lo caratterizzano. Pertanto si ha a disposizione uno strumento molto utile per valutare in che modo cambiano le politiche ottime al variare delle condizioni al contorno. Il prezzo da pagare è naturalmente un alto livello di astrazione.

Un approccio che si avvicina a quello seguito in questa tesi si trova in [20] dove viene considerato un sensore radio alimentato da energia solare. Tuttavia in questo lavoro l'attività del dispositivo viene ridotta ad un'alternanza di fasi in cui il nodo è attivo con altre in cui risulta inattivo; non si considerano cioè le operazioni specifiche compiute dal dispositivo e i relativi consumi energetici. Il sistema è rappresentato mediante una catena di Markov multidimensionale e attraverso la teoria dei giochi vengono studiate strategie ottime per il passaggio del dispositivo dallo stato attivo a quello inattivo. Anche in [27] viene proposto un modello stocastico per un dispositivo dotato di pannello solare, ma l'attenzione è rivolta al problema del dimensionamento delle componenti del dispositivo per supportare una data applicazione. La problematica affrontata in [24] e [22] è invece quella di minimizzare l'energia consumata dal nodo rispettando al contempo le esigenze applicative; vengono sviluppate delle politiche di gestione energetica che tengano conto dello stato di carica della batteria, tuttavia non

è previsto che questa possa essere ricaricata mediante energia prelevata dall'ambiente circostante. Infine è interessante la trattazione contenuta in [15]: il sensore radio è modellato, come verrà fatto in questa tesi, mediante un processo decisionale di Markov a tempo disceto e vengono determinate delle politiche ottime per l'utilizzo efficiente della batteria. Il modello però non tiene conto in alcun modo del livello di carica della batteria e non ammette la presenza di alcuna sorgente di energia che possa ricaricarla.

1.1 Struttura della tesi

Il *capitolo 2* della tesi è dedicato alla presentazione dei risultati teorici su cui si basa la realizzazione del modello del sensore radio alimentato da una fonte di energia rinnovabile; in particolare si fa riferimento alla programmazione dinamica stocastica. Si mostrano la notazione e le definizioni che saranno utilizzate nel resto della trattazione e si introducono alcune assunzioni che consentono di concentrare l'attenzione sulla porzione d'interesse della vasta teoria dei processi decisionali di Markov. Nello specifico si considera il criterio di ottimizzazione del costo medio per slot, discutendo alcuni risultati fondamentali ad esso relativi.

Il *capitolo 3* contiene il contributo principale ed originale di questa tesi: un modello matematico in grado di rappresentare, nella sua complessità, un sensore radio alimentato da una fonte di energia rinnovabile. Si presentano in dettaglio le componenti fondamentali e i parametri che caratterizzano il modello proposto. Si giunge a descrivere il funzionamento del dispositivo mediante un processo decisionale di Markov per il quale valgono i risultati illustrati nel capitolo 2. Il modello realizzato consente di formulare in modo formale e matematico il problema di determinare quale sia il comportamento ottimale per il sensore radio, cioè quali siano le politiche di scheduling ottime che il dispositivo deve seguire al fine di massimizzare le prestazioni.

Nel *capitolo 4* si mettono in evidenza alcune caratteristiche strutturali del modello ideato e presentato nel capitolo 3. In base ad esse si individua una procedura di ottimizzazione che consenta di determinare le politiche di scheduling ottime per il sensore radio. Si introduce, quindi, l'algoritmo di ottimizzazione effettivamente utilizzato ed implementato.

Infine nel *capitolo 5* si discutono alcuni risultati ottenuti utilizzando un ottimizzatore realizzato in C++ che implementa l'algoritmo di ottimizzazione presentato nel capitolo 4 applicandolo al modello proposto nel capitolo 3. Si studiano gli effetti della variazione di alcuni parametri del modello sulle politiche di scheduling ottime calcolate mediante l'ottimizzatore e sulle prestazioni che il dispositivo è in grado di raggiungere seguendo tali politiche.

2

Processi decisionali di Markov

La programmazione dinamica stocastica e in particolare il concetto di processo decisionale di Markov costituiscono le fondamenta teoriche per la trattazione successiva. In questo capitolo si introducono la notazione e le definizioni principali in base alle quali si procederà alla realizzazione del modello del sensore radio. Vengono proposti, inoltre, alcuni risultati fondamentali che saranno ampiamente utilizzati per l'ottimizzazione.

2.1 Introduzione ai processi decisionali di Markov

La teoria dei processi decisionali di Markov, nota anche come *Programmazione Dinamica Stocastica* o *SDP (Stochastic Dynamic Programming)*, fornisce degli strumenti molto utili per il controllo ottimo di un sistema che evolve in modo stocastico.

L'oggetto di studio, dunque, è un sistema il cui comportamento non è prefissato o noto a priori bensì probabilistico. In determinati istanti temporali, inoltre, il sistema esegue delle azioni che influenzano la sua evoluzione futura. L'interesse è quello di definire un'*unità di controllo* o *controller* che sia in grado di decidere in maniera appropriata quale azione il sistema debba compiere di volta in volta. Dunque ci si trova di fronte al problema, o se vogliamo all'opportunità, di condizionare l'evoluzione nel tempo di un sistema probabilistico. L'obiettivo finale della SDP è ottenere un controller capace di scegliere una sequenza di azioni che porti il sistema ad operare in modo ottimo rispetto ad un determinato criterio prestabilito.

L'esperienza quotidiana insegna che decidere in un determinato momento di compiere una certa azione ha un duplice effetto: acquisire un costo (o guadagno) immediato dovuto alla particolare azione scelta e influenzare la dinamica futura e quindi i costi che saranno accumulati successivamente. Optare per l'azione che porta al costo minimo nell'immediato tipicamente non è la soluzione migliore se l'interesse è quello di minimizzare il costo complessivo. Si pensi, ad esempio, ad una maratona: partire alla

massima velocità porta ad esaurire rapidamente le energie e dunque può non essere la scelta migliore se l'obiettivo è terminare la gara nel minor tempo possibile. Il controller, quindi, non deve essere "miope" nel decidere quale sia l'azione ottimale che il sistema deve compiere in un determinato momento per raggiungere un dato obiettivo globale.

Il concetto appena esposto è alla base della SDP: ciascuna azione deve essere valutata in base al suo costo immediato e al costo che si potrebbe accumulare in futuro se l'azione venisse effettivamente scelta. Per tradurre in pratica questa idea generale e piuttosto intuitiva è di cruciale importanza introdurre una definizione precisa di processo decisionale di Markov [1, 2, 21, 25].

2.2 Definizioni e notazione

Si consideri un sistema la cui evoluzione è stocastica e condizionata dalla particolare sequenza di azioni compiute in determinati istanti temporali. Un processo decisionale di Markov consente di rappresentare in modo formale un sistema di questo tipo ed è composto da cinque elementi fondamentali: gli istanti di decisione, gli stati, le azioni, le probabilità di transizione e i costi.

2.2.1 Istanti di decisione

Gli istanti temporali nei quali l'unità di controllo decide quale azione il sistema debba compiere in quel momento si definiscono *istanti di decisione*. L'insieme \mathcal{T} di tutti gli istanti di decisione è un sottoinsieme del semiasse reale non negativo. Può essere un insieme discreto o continuo e, inoltre, può essere finito o infinito.

Particolarmente interessante per la trattazione successiva è il caso in cui il sistema considerato sia a tempo discreto. Il tempo allora è suddiviso e misurato in *slot* o *stage*. In questo frangente l'insieme \mathcal{T} è discreto e si assume che gli istanti decisionali coincidano con l'inizio di ciascuno slot. Inoltre \mathcal{T} può essere finito, quindi del tipo $\mathcal{T} = \{0, 1, 2, \dots, N\}$ per qualche intero positivo $N < \infty$, oppure infinito, cioè $\mathcal{T} = \{0, 1, 2, \dots\}$. Se \mathcal{T} è finito si è soliti parlare di *orizzonte finito*, viceversa quando \mathcal{T} è infinito si parla di *orizzonte infinito*.

2.2.2 Stati

In corrispondenza di ciascun istante di decisione il sistema occupa un determinato *stato*. Lo stato rappresenta in modo completo ed esaustivo la condizione in cui si trova il sistema in un dato momento.

Sia $t \in \mathcal{T}$ un generico istante di decisione. Si indica con \mathcal{S}_t l'insieme di tutti i possibili stati del sistema nell'istante t . L'insieme $\mathcal{S} = \cup_{t \in \mathcal{T}} \mathcal{S}_t$ comprende, quindi, tutti i possibili stati che il sistema può occupare durante la sua evoluzione. Ciascuno degli \mathcal{S}_t e conseguentemente anche \mathcal{S} può essere un insieme finito o infinito. Se \mathcal{S} è finito si parla di sistema, e quindi di processo, a *stati finiti* altrimenti di sistema a *stati infiniti*.

Molto spesso capita di considerare sistemi il cui insieme degli stati è invariante nel tempo. In tal caso $\mathcal{S}_t = \mathcal{S}, \forall t \in \mathcal{T}$.

2.2.3 Azioni

Si consideri un generico istante di decisione $t \in \mathcal{T}$. Il sistema occupa uno degli stati $i \in \mathcal{S}_t$ e l'unità di controllo ha il compito di decidere quale azione deva essere eseguita. La scelta viene fatta tra tutte le possibili azioni che sono consentite nello stato i nell'istante t ; tali azioni costituiscono il cosiddetto *action set* dello stato i nell'istante t indicato con $\mathcal{A}_t(i)$. L'insieme di tutte le azioni che il sistema può compiere nella sua evoluzione è $\mathcal{A} = \cup_{i \in \mathcal{S}_t}^t \mathcal{A}_t(i)$. Ciascuno degli action set $\mathcal{A}_t(i)$ e quindi anche l'insieme \mathcal{A} può essere finito o infinito.

In molte applicazioni sia l'insieme degli stati che gli action set di ciascuno stato risultano invarianti nel tempo. In questo caso $\mathcal{A}_t(i) = \mathcal{A}(i), \forall t \in \mathcal{T}$, e $\mathcal{A} = \cup_{i \in \mathcal{S}} \mathcal{A}(i)$.

2.2.4 Probabilità di transizione

Si assuma che in corrispondenza dell'istante di decisione $t \in \mathcal{T}$ il sistema si trovi nello stato $i \in \mathcal{S}_t$ e esegua l'azione a scelta dal controller tra tutte le possibili azioni in $\mathcal{A}_t(i)$. Lo stato occupato dal sistema nell'istante di decisione successivo è determinato in base alla distribuzione di probabilità $p_t(\cdot|i, a)$ nota come *probabilità di transizione*. In particolare fissato $j \in \mathcal{S}_{t+1}$ il termine $p_t(j|i, a)$ indica la probabilità che il sistema si

sposti nello stato j all'istante $t + 1$ dato che all'istante t lo stato è i e l'azione scelta è a . Si assume che valga la condizione di normalizzazione

$$\sum_{j \in \mathcal{S}_{t+1}} p_t(j|i, a) = 1, \quad t \in \mathcal{T}, i \in \mathcal{S}_t, a \in \mathcal{A}_t(i).$$

Tipicamente anche le probabilità di transizione sono indipendenti da t . Si ha allora $p_t(j|i, a) = p(j|i, a), \forall t \in \mathcal{T}$, e il termine $p(j|i, a)$ può essere indicato in modo più compatto con $p_{ij}(a)$.

2.2.5 Costi

Se lo stato occupato dal sistema all'istante $t \in \mathcal{T}$ è $i \in \mathcal{S}_t$ e viene eseguita l'azione $a \in \mathcal{A}_t(i)$ l'unità di controllo accumula un *costo* che dipende dalla particolare azione scelta e dallo stato; tale costo viene indicato con $g_t(i, a)$. Si assume, inoltre, che la funzione costo $g_t(i, a)$ assuma valori reali non negativi. A seconda del sistema considerato i costi possono essere limitati superiormente o meno.

In alcune applicazioni può succedere che i costi dipendano anche dallo stato j raggiunto all'istante successivo $t + 1$. In questo caso più generale il costo accumulato si indica con $g_t(i, a, j)$ e $g_t(i, a)$ rappresenta invece il costo accumulato mediato su tutti i possibili stati d'arrivo, ovvero:

$$g_t(i, a) = \sum_{j \in \mathcal{S}_{t+1}} g_t(i, a, j) p_t(j|i, a).$$

Anche i costi molto spesso sono indipendenti dall'istante t considerato. In tal caso si ha $g_t(i, a) = g(i, a), \forall t \in \mathcal{T}$.

2.2.6 Processo decisionale di Markov

Quelli appena presentati sono gli ingredienti fondamentali di un processo decisionale di Markov.

Definizione 2.2.1 Si definisce processo decisionale di Markov o MDP (Markov Decision Process) la collezione di oggetti

$$\{\mathcal{T}, \mathcal{S}_t, \mathcal{A}_t(i), p_t(\cdot|i, a), g_t(i, a)\},$$

con $t \in \mathcal{T}$, $i \in \mathcal{S}_t$ e $a \in \mathcal{A}_t(i)$.

L'appellativo "Markov" è utilizzato poiché le probabilità di transizione e i costi dipendono dall'evoluzione passata del sistema solamente attraverso lo stato corrente e l'azione scelta dall'unità di controllo in tale stato.

Si definisce poi *problema decisionale di Markov* (*Markov decision problem*) un processo decisionale di Markov con un determinato criterio di ottimizzazione. Per capire cosa si intende esattamente per criterio di ottimizzazione è necessario introdurre i concetti fondamentali di regola di decisione e politica.

2.2.7 Regola di decisione

Una *regola di decisione* associa ad ogni possibile stato che può essere occupato dal sistema in un determinato istante di decisione, l'azione che deve essere eseguita. In sostanza si tratta di una funzione $\mu_t : \mathcal{S}_t \rightarrow \mathcal{A}_t(i)$ che specifica l'azione da scegliere quando il sistema occupa lo stato $i \in \mathcal{S}_t$ all'istante t ; naturalmente per ogni $i \in \mathcal{S}_t$, $\mu_t(i) \in \mathcal{A}_t(i)$. Quindi fissare una regola di decisione μ_t per un determinato istante t significa indicare quale azione il sistema deva compiere per tutti i possibili stati che il sistema può occupare in quell'istante. L'insieme di tutte le regole di decisione ammissibili all'istante t è indicato con \mathcal{D}_t .

Si noti che una volta definita la regola di decisione μ_t per l'istante t le probabilità di transizione e i costi diventano funzioni del solo stato i occupato dal sistema in t . In particolare il costo accumulato è $g_t(i, \mu_t(i))$ mentre le probabilità di transizione diventano $p_t(j|i, \mu_t(i))$.

2.2.8 Politica

Una *politica* specifica la sequenza di regole di decisione da applicare in ciascun istante di decisione. Essa fornisce all'unità di controllo una strategia completa per la selezione delle azioni che il sistema deve compiere qualsiasi sia la sua effettiva evoluzione. Una politica π è quindi una sequenza di regole di decisione del tipo $\pi = \{\mu_0, \mu_1, \mu_2, \dots\}$ dove $\mu_t \in \mathcal{D}_t$ per $t = 0, 1, 2, \dots$. Si indica con Π l'insieme di tutte le possibili politiche.

Ricoprono un ruolo cruciale le cosiddette *politiche stazionarie*. Una politica si dice stazionaria se $\mu_t = \mu, \forall t \in \mathcal{T}$ ovvero se è del tipo $\pi = \{\mu, \mu, \dots\}$. Adottare una politica stazionaria significa quindi applicare la medesima regola di decisione μ in ciascun istante. Per semplificare la notazione si indica la politica stazionaria $\pi = \{\mu, \mu, \dots\}$ semplicemente con μ .

2.3 Criterio di ottimizzazione

Dopo avere rappresentato un sistema stocastico mediante un processo decisionale di Markov l'obiettivo è quello di individuare una politica che, una volta applicata dall'unità di controllo, porti il sistema ad operare in modo ottimo. Per stabilire quale comportamento del sistema sia ottimo è necessario definire il concetto di ottimalità e, quindi, un criterio di ottimizzazione.

Sono molteplici i criteri di ottimizzazione che possono essere adottati a seconda del tipo di sistema considerato e del risultato che si vuole raggiungere. In base alle caratteristiche del sistema che si andrà a modellare nel capitolo 3 è utile allora introdurre delle assunzioni che consentano di individuare nel vasto mondo della programmazione dinamica stocastica i risultati effettivamente utili nel caso di interesse.

2.3.1 Assunzioni

Di seguito sono elencate le assunzioni che da questo momento in poi si supporranno soddisfatte dal sistema considerato.

- *Sistema a tempo discreto*: il tempo è suddiviso e misurato in slot. Gli istanti di decisione corrispondono all'inizio di ciascuno slot. L'azione scelta in un istante di decisione viene svolta durante lo slot e porta il sistema in un nuovo stato all'istante successivo in base alle corrispondenti probabilità di transizione.
- *Orizzonte infinito*: l'insieme $\mathcal{T} = \{0, 1, 2, \dots\}$ degli istanti di decisione è infinito. L'evoluzione del sistema non ha dunque una durata finita.
- *Sistema stazionario*: le caratteristiche strutturali del sistema sono invarianti nel tempo. L'insieme degli stati, le azioni possibili in ciascuno stato, le probabilità

di transizione e i costi sono indipendenti dall'istante $t \in \mathcal{T}$ considerato. In ogni istante di decisione il sistema si trova in un qualche stato $i \in \mathcal{S}$, sceglie di eseguire un'azione $a \in \mathcal{A}(i)$ accumulando un costo $g(i, a)$ e si sposta nello stato $j \in \mathcal{S}$ all'istante successivo con probabilità $p_{ij}(a)$.

- *Insieme degli stati finito*: L'insieme \mathcal{S} di tutti i possibili stati che il sistema può occupare in un dato istante è finito. Si indica con S la cardinalità di \mathcal{S} , $S = |\mathcal{S}|$.
- *Action set finiti*: per ogni stato $i \in \mathcal{S}$ l'action set $\mathcal{A}(i)$ è finito. Naturalmente anche $\mathcal{A} = \cup_{i \in \mathcal{S}} \mathcal{A}(i)$ è un insieme finito.
- *Costi limitati*: il costo accumulato in un qualsiasi slot è superiormente limitato. Esiste, cioè, un intero positivo $C < \infty$ tale che $g(i, a) < C$ per ogni $i \in \mathcal{S}$ e $a \in \mathcal{A}(i)$.
- *Conoscenza completa*: la politica che l'unità di controllo deve applicare al sistema viene determinata assumendo che il controller abbia in ciascun istante di decisione una conoscenza completa dello stato corrente occupato dal sistema.

In base a queste assunzioni è possibile introdurre il criterio di ottimizzazione a cui si farà riferimento nel prosieguo.

2.3.2 Costo medio per slot

Si considera un processo decisionale di Markov mediante il quale è rappresentato un sistema che soddisfa le assunzioni di sezione 2.3.1. Si indica con $\pi = \{\mu_0, \mu_1, \dots\}$ una generica politica applicata dall'unità di controllo al sistema; naturalmente $\mu_t(i) \in \mathcal{A}(i)$ per ogni i e t . Con s_t invece è indicato lo stato occupato dal sistema nell'istante $t \in \mathcal{T}$; lo stato s_t è aleatorio e assume i valori in \mathcal{S} con probabilità che dipendono dalla particolare politica π scelta e dalle probabilità di transizione.

Fissato lo stato iniziale $s_0 = i \in \mathcal{S}$, si definisce *costo medio per slot* (*average cost per slot*) della politica π a partire dallo stato i la quantità

$$J_\pi(i) = \limsup_{N \rightarrow \infty} \frac{1}{N} E \left[\sum_{t=0}^{N-1} g(s_t, \mu_t(s_t)) \middle| s_0 = i \right]. \quad (2.1)$$

L'utilizzo del limite superiore è dettato dal fatto che esso esiste certamente per ogni stato iniziale i e per ogni politica π , cosa non necessariamente vera per il limite semplice.

Si definisce costo medio per slot *ottimo* a partire dallo stato i il valore minimo del costo medio per slot valutato su tutte le politiche ammissibili, ovvero la quantità

$$J(i) = \min_{\pi \in \Pi} J_{\pi}(i). \quad (2.2)$$

Si può introdurre a questo punto la seguente definizione di fondamentale importanza.

Definizione 2.3.1 *Una politica $\pi \in \Pi$ ammissibile per il sistema considerato si dice ottima secondo il criterio del costo medio per slot se e solo se*

$$J_{\pi}(i) = J(i), \quad i \in \mathcal{S}.$$

Secondo questo criterio di ottimizzazione, quindi, una politica è ottima se porta il sistema a minimizzare il costo medio accumulato per slot.

Molti dei risultati teorici relativi al criterio del costo medio per slot che si vedranno in sezione 2.4 si basano sulla relazione che intercorre tra questo criterio e un altro criterio di ottimizzazione che ora viene introdotto.

2.3.3 Costo totale α -scontato

Come in sezione 2.3.2, si considera un sistema per cui valgono le assunzioni di sezione 2.3.1 e si indica con $\pi = \{\mu_0, \mu_1, \dots\}$ una generica politica applicata dall'unità di controllo al sistema.

Fissati lo stato iniziale $s_0 = i \in \mathcal{S}$ e la costante $\alpha \in (0, 1)$, si definisce *costo totale atteso α -scontato* (*expected total α -discounted cost*) della politica π a partire dallo stato i la quantità

$$J_{\alpha, \pi}(i) = \lim_{N \rightarrow \infty} E \left[\sum_{t=0}^{N-1} \alpha^t g(s_t, \mu_t(s_t)) \middle| s_0 = i \right]. \quad (2.3)$$

La costante α specifica il fattore di sconto applicato ai costi futuri. L'idea, propria dell'ambito economico, è che i costi accumulati nell'evoluzione futura abbiano un valore scontato secondo la valuta corrente.

Il costo totale α -scontato *ottimo* a partire dallo stato i è dunque

$$J_\alpha(i) = \min_{\pi \in \Pi} J_{\alpha,\pi}(i). \quad (2.4)$$

Da ciò deriva la definizione seguente.

Definizione 2.3.2 Una politica $\pi \in \Pi$ ammissibile per il sistema considerato si dice ottima secondo il criterio del costo totale α -scontato se e solo se

$$J_{\alpha,\pi}(i) = J_\alpha(i), \quad i \in \mathcal{S}.$$

In questo caso, quindi, una politica è ottima se porta il sistema a minimizzare il costo atteso complessivo con fattore di sconto α . Naturalmente questo criterio di ottimizzazione può essere adottato quando è sensato introdurre un fattore di sconto sui costi futuri.

2.4 Ottimizzazione secondo il criterio del costo medio per slot

Si concentra l'attenzione su un problema decisionale di Markov in cui il criterio di ottimizzazione considerato è quello del costo medio per slot e in cui si suppongono soddisfatte le assunzioni di sezione 2.3.1. Si vuole determinare una politica che risulti ottima in base alla definizione 2.3.1 e che quindi, una volta applicata al sistema dal controller, renda minimo il costo medio per slot accumulato [2].

In questo contesto ricoprono un ruolo fondamentale le politiche stazionarie. Data una generica politica stazionaria μ , si indica con $J_\mu(i)$ il costo medio per slot di μ a partire dallo stato i e si utilizza la seguente notazione compatta dove $\mathcal{S} = \{1, 2, \dots, S\}$ è l'insieme degli stati del sistema:

$$g_\mu = \begin{bmatrix} g(1, \mu(1)) \\ \vdots \\ g(S, \mu(S)) \end{bmatrix}, \quad P_\mu = \begin{bmatrix} p_{11}(\mu(1)) & \cdots & p_{1S}(\mu(1)) \\ \vdots & \ddots & \vdots \\ p_{S1}(\mu(S)) & \cdots & p_{SS}(\mu(S)) \end{bmatrix}, \quad J_\mu = \begin{bmatrix} J_\mu(1) \\ \vdots \\ J_\mu(S) \end{bmatrix}.$$

È utile osservare che una volta fissata una politica stazionaria μ essa induce nel processo decisionale una catena di Markov a stati finiti dove lo spazio degli stati è \mathcal{S} e

la matrice delle probabilità di transizione è P_μ . In base alle caratteristiche strutturali delle catene di Markov indotte dalle politiche stazionarie diversi sono i risultati teorici e gli algoritmi che possono essere adottati per l'ottimizzazione del costo medio per slot. Alcuni concetti basilari sulle catene di Markov sono pertanto di fondamentale importanza e vengono di seguito richiamati.

2.4.1 Definizioni utili sulle catene di Markov a stati finiti

La teoria delle catene di Markov a stati finiti gioca un ruolo cruciale nell'ottimizzazione del costo medio per slot. Di seguito vengono ricordate solo alcune definizioni utili per la trattazione successiva. Per un'analisi più approfondita si rimanda a [3, 5, 13].

Si consideri una catena di Markov a stati finiti con matrice delle probabilità di transizione P . Uno stato j si dice *raggiungibile* dallo stato i se è positiva la probabilità di raggiungere lo stato j a partire dallo stato i in un qualche numero finito di passi. Due stati i e j si definiscono *comunicanti* se sono raggiungibili l'uno dall'altro. La comunicazione è una relazione di equivalenza e in quanto tale partiziona l'insieme degli stati in classi di equivalenza dette classi comunicanti. Una *classe comunicante* è costituita da tutti e soli gli stati che comunicano tra loro.

Uno stato i si dice *ricorrente* se e solo se, una volta che la catena è partita dallo stato i , la probabilità di ritorno nello stato i in un qualche numero finito di transizioni è uno. Pertanto uno stato ricorrente, dopo essere stato visitato una volta dalla catena, viene visitato un numero infinito di volte con probabilità uno. Uno stato i si dice *transitorio* se non è ricorrente. La ricorrenza è una proprietà di classe, cioè se uno stato è ricorrente lo sono anche tutti gli stati appartenenti alla medesima classe comunicante detta, in tal caso, *classe ricorrente*. Da un generico stato appartenente ad una classe ricorrente vi è probabilità uno di raggiungere un qualsiasi altro stato della classe e probabilità zero di raggiungere uno stato all'esterno di essa. Si può dimostrare [5] che ogni catena di Markov con un numero finito di stati possiede almeno una classe ricorrente.

Si definisce *periodo* di uno stato i il massimo comune divisore di tutti gli interi $k \geq 1$ tali per cui la probabilità di transitare dallo stato i a i stesso in k passi è positiva. Se lo stato i ha periodo $d = 1$ si dice *aperiodico*, se ha periodo $d > 1$ si dice *periodico* di

periodo d . Anche la periodicità è una proprietà di classe; una *classe periodica* di periodo d è costituita tutta da stati periodici con la medesima periodicità. Una catena di Markov o equivalentemente la sua matrice delle probabilità di transizione P si dice periodica se comprende almeno una classe periodica, in caso contrario si dice aperiodica.

Si dimostra [3, 13] che una catena di Markov con matrice P è aperiodica se e solo se vale la relazione

$$\lim_{t \rightarrow \infty} P^t = P^*, \quad (2.5)$$

dove la matrice P^* è definita come

$$P^* = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} P^t. \quad (2.6)$$

Il limite di equazione (2.6) esiste sempre qualsiasi sia la catena di Markov P considerata. La componente (i, j) -esima della matrice P^* , $[P^*]_{ij}$, rappresenta la frequenza relativa a lungo termine con la quale viene visitato lo stato j dato che lo stato iniziale è i . Pertanto se i è uno stato ricorrente $[P^*]_{ij} > 0$ se e solo se j è ricorrente e appartiene alla stessa classe ricorrente di i . Inoltre $[P^*]_{ij} = 0$ per ogni i se e solo se j è uno stato transitorio.

2.4.2 Relazione tra costo medio per slot e costo totale α -scontato

Sia μ una generica politica stazionaria e sia $J_{\alpha, \mu}(i)$ il suo costo totale α -scontato a partire dallo stato i , con $\alpha \in (0, 1)$. Posto

$$J_{\alpha, \mu} = \begin{bmatrix} J_{\alpha, \mu}(1) \\ \vdots \\ J_{\alpha, \mu}(S) \end{bmatrix},$$

vale l'equazione

$$J_{\alpha, \mu} = \lim_{N \rightarrow \infty} \sum_{t=0}^{N-1} \alpha^t P_{\mu}^t g_{\mu} = \left(\sum_{t=0}^{\infty} \alpha^t P_{\mu}^t \right) g_{\mu} = (I - \alpha P_{\mu})^{-1} g_{\mu}. \quad (2.7)$$

Molto importante è la relazione che intercorre tra J_μ e $J_{\alpha,\mu}$. Prima di mostrarla formalmente è utile introdurla in modo intuitivo. Considerato un generico stato $i \in \mathcal{S}$, si può scrivere

$$\begin{aligned} J_\mu(i) &= \limsup_{N \rightarrow \infty} \frac{1}{N} E \left[\sum_{t=0}^{N-1} g(s_t, \mu(s_t)) \mid s_0 = i \right] \\ &= \limsup_{N \rightarrow \infty} \lim_{\alpha \rightarrow 1} \frac{E \left[\sum_{t=0}^{N-1} \alpha^t g(s_t, \mu(s_t)) \mid s_0 = i \right]}{\sum_{t=0}^{N-1} \alpha^t}. \end{aligned}$$

Assumendo che l'ordine dei limiti possa essere invertito si ottiene

$$\begin{aligned} J_\mu(i) &= \lim_{\alpha \rightarrow 1} \limsup_{N \rightarrow \infty} \frac{E \left[\sum_{t=0}^{N-1} \alpha^t g(s_t, \mu(s_t)) \mid s_0 = i \right]}{\sum_{t=0}^{N-1} \alpha^t} \\ &= \lim_{\alpha \rightarrow 1} \frac{\lim_{N \rightarrow \infty} E \left[\sum_{t=0}^{N-1} \alpha^t g(s_t, \mu(s_t)) \mid s_0 = i \right]}{\lim_{N \rightarrow \infty} \sum_{t=0}^{N-1} \alpha^t} \\ &= \lim_{\alpha \rightarrow 1} (1 - \alpha) J_{\alpha,\mu}(i). \end{aligned}$$

Dato che il procedimento è valido per ogni $i \in \mathcal{S}$ si può concludere che

$$J_\mu = \lim_{\alpha \rightarrow 1} (1 - \alpha) J_{\alpha,\mu}. \quad (2.8)$$

Per la dimostrazione rigorosa della relazione (2.8) è fondamentale il seguente teorema.

Teorema 2.4.1 *Per ogni matrice stocastica P e per ogni $\alpha \in (0, 1)$ vale l'equazione*

$$(I - \alpha P)^{-1} = (1 - \alpha)^{-1} P^* + H + O(|1 - \alpha|), \quad (2.9)$$

dove $O(|1 - \alpha|)$ è una matrice dipendente da α tale che

$$\lim_{\alpha \rightarrow 1} O(|1 - \alpha|) = 0,$$

mentre le matrici P^* e H sono date da

$$P^* = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} P^t, \quad (2.10)$$

$$H = (I - P + P^*)^{-1} - P^*. \quad (2.11)$$

Inoltre le matrici P^* e H soddisfano le seguenti equazioni:

$$P^* = PP^* = P^*P = P^*P^*, \quad (2.12)$$

$$P^*H = HP^* = 0, \quad (2.13)$$

$$P^* + H = I + PH. \quad (2.14)$$

Dimostrazione. La regola di Cramer per l'inversione di una matrice porta ad esprimere ciascun elemento dell'inversa di una matrice come rapporto tra due determinanti. Definita allora la matrice

$$M(\alpha) = (1 - \alpha)(I - \alpha P)^{-1},$$

i suoi elementi sono nulli oppure frazioni in cui il numeratore e il denominatore sono polinomi in α privi di divisori comuni. I polinomi al denominatore degli elementi non nulli di $M(\alpha)$ non possono avere 1 come radice. Se così fosse vi sarebbero degli elementi di $M(\alpha)$ che tendono all'infinito per $\alpha \rightarrow 1$; ma questo non è possibile. Infatti, a partire dall'equazione (2.7) valida per una qualsiasi μ , si può scrivere $(1 - \alpha)^{-1}M(\alpha)g_\mu = (I - \alpha P)^{-1}g_\mu = J_{\alpha,\mu}$ e $J_{\alpha,\mu}(i) \leq (1 - \alpha)^{-1} \max_j \{g_\mu(j)\}$, $i \in \mathcal{S}$. Questo implica che gli elementi di $M(\alpha)g_\mu$ sono limitati da $\max_j \{g_\mu(j)\}$ per ogni $\alpha < 1$; ma allora lo sono anche gli elementi di $M(\alpha)$. Si può concludere quindi che l' (i, j) -esimo elemento di $M(\alpha)$ è del tipo

$$[M(\alpha)]_{ij} = \frac{\lambda(\alpha - z_1) \cdots (\alpha - z_q)}{(\alpha - r_1) \cdots (\alpha - r_p)},$$

dove $\lambda, z_i, i = 1, \dots, q$, e $r_i, i = 1, \dots, p$, sono scalari tali che $r_i \neq 1$ per $i = 1, \dots, p$.

Sia ora

$$P^* = \lim_{\alpha \rightarrow 1} M(\alpha), \quad (2.15)$$

e sia H la matrice il cui elemento (i, j) -esimo è la derivata prima di $-[M(\alpha)]_{ij}$ valutata in $\alpha = 1$. Considerata l'espansione in serie di Taylor troncata al primo ordine di $[M(\alpha)]_{ij}$, per ogni α nell'intorno di $\alpha = 1$ si ha

$$(1 - \alpha)(I - \alpha P)^{-1} = M(\alpha) = P^* + (1 - \alpha)H + O((1 - \alpha)^2), \quad (2.16)$$

dove $O((1 - \alpha)^2)$ è una matrice dipendente da α tale che

$$\lim_{\alpha \rightarrow 1} \frac{O((1 - \alpha)^2)}{(1 - \alpha)} = 0.$$

Moltiplicando entrambi i membri dell'equazione (2.16) per $(1 - \alpha)^{-1}$ si ottiene la relazione (2.9) voluta. Naturalmente rimane da dimostrare che le matrici P^* e H appena definite soddisfano anche le equazioni (2.10) e (2.11).

Presa in considerazione la matrice P^* così come definita in (2.15), si dimostra a questo punto che essa soddisfa nell'ordine le equazioni (2.12), (2.11), (2.13), (2.14) e (2.10).

Innanzitutto è immediato osservare che vale la relazione

$$(I - \alpha)(I - \alpha P)(I - \alpha P)^{-1} = (1 - \alpha)I. \quad (2.17)$$

Dall'equazione (2.17) segue immediatamente

$$\alpha P(1 - \alpha)(I - \alpha P)^{-1} = (1 - \alpha)(I - \alpha P)^{-1} + (\alpha - 1)I.$$

Eseguendo il limite per $\alpha \rightarrow 1$ e ricordando la definizione (2.15) si ottiene

$$PP^* = P^*.$$

Se poi nell'equazione (2.17) si inverte l'ordine di $(I - \alpha P)$ e $(I - \alpha P)^{-1}$ e si ripete il medesimo ragionamento si ricava che $P^*P = P^*$. Da $PP^* = P^*$ si ottiene anche $(I - \alpha P)P^* = (1 - \alpha)P^*$, da cui $P^* = (1 - \alpha)(I - \alpha P)^{-1}P^*$. Prendendo il limite per $\alpha \rightarrow 1$ e utilizzando ancora la definizione (2.15) si ottiene $P^* = P^*P^*$. L'equazione (2.12) è quindi dimostrata.

Sfruttando la relazione (2.12) si può scrivere $(P - P^*)^2 = P^2 - P^*$ e, procedendo per induzione, si ottiene

$$(P - P^*)^t = P^t - P^*, \quad t = 1, 2, 3, \dots$$

Pertanto

$$\begin{aligned} (I - \alpha P)^{-1} - (1 - \alpha)^{-1}P^* &= \sum_{t=0}^{\infty} \alpha^t (P^t - P^*) \\ &= I - P^* + \sum_{t=1}^{\infty} \alpha^t (P - P^*)^t \\ &= (I - \alpha(P - P^*))^{-1} - P^*. \end{aligned}$$

D'altra parte dall'equazione (2.16) si ricava

$$\begin{aligned} H &= \lim_{\alpha \rightarrow 1} ((1 - \alpha)^{-1} M(\alpha) - (1 - \alpha)^{-1} P^*) \\ &= \lim_{\alpha \rightarrow 1} ((I - \alpha P)^{-1} - (1 - \alpha)^{-1} P^*). \end{aligned}$$

Combinando le due relazioni appena ottenute si ha

$$H + P^* = \lim_{k \rightarrow \infty} A_k, \quad (2.18)$$

dove

$$A_k = (I - \alpha_k(P - P^*))^{-1}$$

e $\{\alpha_k\}$ è una sequenza tale che $\alpha_k \rightarrow 1$ per $k \rightarrow \infty$. Moltiplicando il membro sinistro dell'equazione (2.18) per $(I - P + P^*)$ e il membro destro per l'equivalente matrice $\lim_{k \rightarrow \infty} A_k^{-1}$ si ottiene

$$(I - P + P^*)(H + P^*) = \lim_{k \rightarrow \infty} A_k^{-1} \lim_{k \rightarrow \infty} A_k = \lim_{k \rightarrow \infty} (A_k^{-1} A_k) = I,$$

dove la seconda uguaglianza è giustificata dal fatto che i due limiti a sinistra esistono sicuramente. Ma allora

$$H + P^* = (I - P + P^*)^{-1},$$

e la relazione (2.11) è dunque valida.

Dall'equazione (2.11) deriva

$$(I - P + P^*)H = I - (I - P + P^*)P^*$$

che, utilizzando l'equazione (2.12), diventa

$$(I - P + P^*)H = I - P^*. \quad (2.19)$$

Premoltiplicando entrambi i membri di (2.19) per P^* e ricordando ancora (2.12) si ottiene direttamente $P^*H = 0$ che è la prima parte dell'equazione (2.13). Postmoltiplicando entrambi i membri di (2.19) per P^* e utilizzando sempre (2.12) si ricava, invece, $(I - P + P^*)HP^* = 0$ che, data l'invertibilità di $(I - P + P^*)$ implica che $HP^* = 0$. L'equazione (2.13) è quindi dimostrata.

Dalle relazioni (2.19) e (2.13) si ottiene immediatamente

$$P^* + H = I + (P - P^*)H = I + PH - P^*H = I + PH,$$

che è esattamente l'equazione (2.14).

A questo punto rimane da dimostrare solo (2.10). Moltiplicando l'equazione (2.14) per P^t per un qualsiasi $t = 0, 1, \dots$ e ricordando ancora una volta la relazione (2.12) si ottiene

$$P^* + P^t H = P^t + P^{t+1} H, \quad t = 0, 1, 2, \dots$$

Sommando per tutti i $t = 0, 1, \dots, N-1$ si ricava

$$NP^* + H = \sum_{t=0}^{N-1} P^t + P^N H,$$

che, dividendo per N , diventa

$$P^* + \frac{H}{N} = \frac{1}{N} \sum_{t=0}^{N-1} P^t + \frac{P^N H}{N}.$$

Eseguendo il limite per $N \rightarrow \infty$ e tenendo conto del fatto che $H/N \rightarrow 0$ e $P^N/N \rightarrow 0$, dato che P e quindi P^N sono matrici stocastiche, si ottiene l'equazione (2.10). ■

La seguente proposizione è diretta conseguenza del teorema 2.4.1. Essa stabilisce in modo preciso la relazione che intercorre tra il costo medio per slot e il costo totale α -scontato di una qualsiasi politica stazionaria.

Proposizione 2.4.2 *Per ogni politica stazionaria μ e per ogni $\alpha \in (0, 1)$ si ha*

$$J_{\alpha, \mu} = (1 - \alpha)^{-1} J_{\mu} + h_{\mu} + O(|1 - \alpha|), \quad (2.20)$$

dove J_{μ} e h_{μ} sono dati da

$$J_{\mu} = P_{\mu}^* g_{\mu}, \quad h_{\mu} = H_{\mu} g_{\mu}, \quad (2.21)$$

con

$$P_{\mu}^* = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} P_{\mu}^t, \quad H_{\mu} = (I - P_{\mu} + P_{\mu}^*)^{-1} - P_{\mu}^*.$$

Inoltre, sono valide le equazioni

$$J_{\mu} = P_{\mu} J_{\mu}, \quad (2.22)$$

$$J_{\mu} + h_{\mu} = g_{\mu} + P_{\mu} h_{\mu}. \quad (2.23)$$

Dimostrazione. Ogni politica stazionaria μ induce una catena di Markov a stati finiti con matrice delle probabilità di transizione P_μ . In quanto tale, P_μ è una matrice stocastica e dunque per essa valgono i risultati del teorema 2.4.1 per ogni $\alpha \in (0, 1)$. In particolare si ha

$$(I - \alpha P_\mu)^{-1} = (1 - \alpha)^{-1} P_\mu^* + H_\mu + O(|1 - \alpha|), \quad (2.24)$$

con

$$P_\mu^* = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} P_\mu^t, \quad H_\mu = (I - P_\mu + P_\mu^*)^{-1} - P_\mu^*.$$

Si può inoltre scrivere

$$\begin{aligned} J_\mu &= \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} P_\mu^t g_\mu = \left(\limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} P_\mu^t \right) g_\mu \\ &= \left(\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=0}^{N-1} P_\mu^t \right) g_\mu \\ &= P_\mu^* g_\mu, \end{aligned}$$

dove il passaggio dal limite superiore al limite è possibile dato che per il teorema 2.4.1 il limite esiste. Si ottiene quindi per J_μ l'espressione in (2.21). Ponendo allora $h_\mu = H_\mu g_\mu$, sempre come in (2.21), moltiplicando entrambi i membri di (2.24) per g_μ e ricordando la relazione (2.7) si ricava direttamente l'equazione (2.20) che risulta così dimostrata.

Applicando l'equazione (2.14) a P_μ si ottiene

$$P_\mu^* + H_\mu = I + P_\mu H_\mu.$$

Moltiplicando entrambi i membri per g_μ ed usando (2.21) si ricava l'equazione (2.23).

Infine, dato che per la relazione (2.12) si ha $P_\mu^* = P_\mu P_\mu^*$, moltiplicando per P_μ entrambi i membri dell'equazione $J_\mu = P_\mu^* g_\mu$ si ottiene l'espressione (2.22). ■

L'equazione (2.20) è nota come *espansione in serie di Laurent* del costo totale α -scontato di una politica stazionaria μ . I vettori J_μ e h_μ sono univocamente definiti e vengono anche chiamati rispettivamente *gain* e *bias* di μ .

Si noti che con la proposizione 2.4.2 e in particolare con la relazione (2.20) viene dimostrata in modo formale la validità dell'equazione (2.8) che si era precedentemente introdotta in modo intuitivo.

Molto interessanti sono due risultati relativi al bias h_μ . Innanzitutto dalla definizione $h_\mu = H_\mu g_\mu$ e dal fatto che $P_\mu^* H_\mu = 0$ per la relazione 2.13 segue l'equazione

$$P_\mu^* h_\mu = 0. \quad (2.25)$$

In secondo luogo, si assuma che valga la condizione

$$P_\mu^* = \lim_{N \rightarrow \infty} P_\mu^N, \quad (2.26)$$

che, per quanto detto in sezione 2.4.1, è soddisfatta se P_μ corrisponde ad una catena di Markov aperiodica. Tenendo conto che per l'equazione (2.23) si ha $g_\mu - J_\mu = h_\mu - P_\mu h_\mu$, si può scrivere

$$\sum_{t=0}^N P_\mu^t (g_\mu - J_\mu) = \sum_{t=0}^N P_\mu^t (h_\mu - P_\mu h_\mu) = \frac{(I - P_\mu^{N+1})(I - P_\mu)h_\mu}{I - P_\mu} = h_\mu - P_\mu^{N+1} h_\mu.$$

Eseguendo il limite per $N \rightarrow \infty$ e utilizzando le equazioni (2.25) e (2.26) si ottiene

$$h_\mu = \lim_{N \rightarrow \infty} \sum_{t=0}^N P_\mu^t (g_\mu - J_\mu). \quad (2.27)$$

L'equazione (2.27) consente di interpretare il bias h_μ come un *costo relativo*: è la differenza tra il costo totale di μ e il costo totale che si accumulerebbe se i costi sul singolo slot fossero quelli medi J_μ .

2.4.3 Politica Blackwell ottima

L'espansione in serie di Laurent (2.20) evidenzia che il vettore J_μ dei costi medi per slot di una politica stazionaria μ consiste di tre termini dipendenti da α :

$$J_\mu = (1 - \alpha)J_{\alpha,\mu} - (1 - \alpha)h_\mu + O(|1 - \alpha|^2). \quad (2.28)$$

Il termine $(1 - \alpha)J_{\alpha,\mu}$ è dominante per $\alpha \approx 1$ dato che i componenti di $J_{\alpha,\mu}$ tendono a diventare infiniti per $\alpha \rightarrow 1$. Questo porta a pensare che se una politica stazionaria minimizza $J_{\alpha,\mu}$ per tutti gli $\alpha \approx 1$ allora minimizza anche J_μ . Questa osservazione motiva l'introduzione di una classe di politiche stazionarie molto importante dal punto di vista teorico [4].

Definizione 2.4.1 Una politica stazionaria μ si dice Blackwell ottima se è simultaneamente ottima secondo il criterio del costo totale α -scontato per tutti i problemi decisionali con $\alpha \in (\bar{\alpha}, 1)$, dove $\bar{\alpha}$ è un qualche scalare tale che $0 < \bar{\alpha} < 1$.

Secondo l'equazione (2.8) per ogni politica stazionaria μ vale

$$J_\mu = \lim_{\alpha \rightarrow 1} (1 - \alpha) J_{\alpha, \mu}.$$

Poiché una politica stazionaria Blackwell ottima minimizza $J_{\alpha, \mu}$ rispetto a μ per tutti gli α sufficientemente vicini a 1, è facile capire che una politica Blackwell ottima è ottima anche secondo il criterio del costo medio per slot all'interno della classe delle politiche stazionarie. Tuttavia la cosa interessante che si giungerà a dimostrare è che in realtà una politica Blackwell ottima è ottima rispetto a tutte le politiche ammissibili, siano esse stazionarie o no.

Per dimostrare alcuni risultati importanti relativi alle politiche Blackwell ottime è necessario comprendere cosa implichi per una politica stazionaria μ essere ottima secondo il criterio del costo totale α -scontato. Si consideri, quindi, un processo decisionale in cui viene adottato tale criterio di ottimizzazione per $\alpha \in (0, 1)$; si dimostra [2] che il costo totale α -scontato ottimo a partire dallo stato i , indicato con $J_\alpha(i)$, soddisfa l'equazione di ottimalità

$$J_\alpha(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \alpha \sum_{j \in \mathcal{S}} p_{ij}(a) J_\alpha(j) \right], \quad i \in \mathcal{S}. \quad (2.29)$$

Una politica stazionaria μ è ottima secondo il criterio del costo totale α -scontato se e solo se $\mu(i)$ realizza il minimo in equazione (2.29) per ogni $i \in \mathcal{S}$.

Tenendo conto di questo risultato, si introduce, ora, una proposizione che assicura l'esistenza di una politica Blackwell ottima.

Proposizione 2.4.3 Preso in considerazione un qualsiasi problema decisionale, esiste sempre una politica Blackwell ottima.

Dimostrazione. Per ogni politica stazionaria μ vale la relazione (2.7):

$$J_{\alpha, \mu} = (I - \alpha P_\mu)^{-1} g_\mu.$$

In base alla regola di Cramer sull'inversione di una matrice, per ogni μ e per ogni stato i , $J_{\alpha,\mu}(i)$ è una funzione razionale in α ovvero un rapporto tra due polinomi in α . Considerate allora due politiche stazionarie μ e μ' , le curve $J_{\alpha,\mu}(i)$ e $J_{\alpha,\mu'}(i)$ coincidono oppure si intersecano un numero finito di volte nell'intervallo $(0, 1)$ su cui varia α . Dato che esiste solo un numero finito di politiche stazionarie, si può concludere che per ogni stato i esiste una politica stazionaria μ^i e uno scalare $\bar{\alpha}_i \in (0, 1)$ tali che μ^i è ottima secondo il criterio del costo totale α -scontato per $\alpha \in (\bar{\alpha}_i, 1)$ quando lo stato iniziale è i . Ma allora, per ogni i , $\mu^i(i)$ realizza il minimo nell'equazione di ottimalità (2.29) per ogni $\alpha \in (\max_i \bar{\alpha}_i, 1)$.

Si considera a questo punto la politica stazionaria μ^* definita per ogni i come $\mu^*(i) = \mu^i(i)$. Allora, $\mu^*(i)$ realizza il minimo nell'equazione (2.29) per ogni i e per ogni $\alpha \in (\max_i \bar{\alpha}_i, 1)$. Conseguentemente μ^* è ottima secondo il criterio del costo totale α -scontato per ogni $\alpha \in (\max_i \bar{\alpha}_i, 1)$ ed è quindi Blackwell ottima. ■

Le politiche Blackwell ottime costituiscono uno strumento analitico fondamentale per giungere alla formulazione di un'equazione di ottimalità analoga a (2.29) anche per il criterio del costo medio per slot. Il seguente teorema introduce un primo risultato utile in questo senso.

Teorema 2.4.4 *Tutte le politiche Blackwell ottime hanno gli stessi vettori gain e bias. Considerate, cioè, due qualsiasi politiche Blackwell ottime μ e μ' valgono le relazioni*

$$J_\mu = J_{\mu'}, \quad h_\mu = h_{\mu'}. \quad (2.30)$$

Sia allora (J^, h^*) la coppia gain-bias comune a tutte le politiche Blackwell ottime. Sono soddisfatte le equazioni*

$$J^*(i) = \min_{a \in \mathcal{A}(i)} \sum_{j \in \mathcal{S}} p_{ij}(a) J^*(j), \quad i \in \mathcal{S}, \quad (2.31)$$

e

$$J^*(i) + h^*(i) = \min_{a \in \hat{\mathcal{A}}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h^*(j) \right], \quad i \in \mathcal{S}, \quad (2.32)$$

dove, per ogni i , l'insieme $\hat{\mathcal{A}}(i)$ contiene le azioni di $\mathcal{A}(i)$ che realizzano il minimo in (2.31). Inoltre, se μ^* è una politica Blackwell ottima, $\mu^*(i)$ minimizza la parte destra di entrambe le equazioni per ogni i .

Dimostrazione. Considerate due qualsiasi politiche stazionarie Blackwell ottime μ e μ' , in base alla proposizione 2.4.2 si può scrivere

$$\begin{aligned} J_{\alpha,\mu} &= (1 - \alpha)^{-1} J_{\mu} + h_{\mu} + O(|1 - \alpha|), \\ J_{\alpha,\mu'} &= (1 - \alpha)^{-1} J_{\mu'} + h_{\mu'} + O(|1 - \alpha|). \end{aligned}$$

Ma per definizione di politica Blackwell ottima $J_{\alpha,\mu} = J_{\alpha,\mu'}$ per tutti i valori di α sufficientemente vicini a 1. Dunque, eseguendo il limite per $\alpha \rightarrow 1$ nelle due equazioni, si ottiene direttamente $J_{\mu} = J_{\mu'}$ e $h_{\mu} = h_{\mu'}$. Le equazioni (2.30) sono quindi verificate.

Sia ora μ^* una politica Blackwell ottima con coppia gain-bias (J^*, h^*) . Per definizione μ^* è ottima secondo il criterio del costo totale α -scontato per tutti gli $\alpha \in (\bar{\alpha}, 1)$. Ricordando l'equazione di ottimalità (2.29), per ogni politica stazionaria μ e $\alpha \in (\bar{\alpha}, 1)$, deve valere

$$g_{\mu^*} + \alpha P_{\mu^*} J_{\alpha,\mu^*} \leq g_{\mu} + \alpha P_{\mu} J_{\alpha,\mu^*}. \quad (2.33)$$

Inoltre, richiamando ancora la proposizione 2.4.2, per tutti gli $\alpha \in (\bar{\alpha}, 1)$ si ha

$$J_{\alpha,\mu^*} = (1 - \alpha)^{-1} J^* + h^* + O(|1 - \alpha|).$$

Sostituendo l'espansione in serie di Laurent in (2.33) si ottiene la disuguaglianza

$$0 \leq g_{\mu} - g_{\mu^*} + \alpha(P_{\mu} - P_{\mu^*})((1 - \alpha)^{-1} J^* + h^* + O(|1 - \alpha|)), \quad (2.34)$$

che, moltiplicata per $(1 - \alpha)$, diventa

$$0 \leq (1 - \alpha)(g_{\mu} - g_{\mu^*}) + \alpha(P_{\mu} - P_{\mu^*})(J^* + (1 - \alpha)h^* + O((1 - \alpha)^2)).$$

Eseguendo il limite per $\alpha \rightarrow 1$ si ricava $P_{\mu^*} J^* \leq P_{\mu} J^*$; ma per la relazione (2.22) $J^* = P_{\mu^*} J^*$ e dunque $J^* = P_{\mu^*} J^* \leq P_{\mu} J^*$. Questa relazione è equivalente a (2.31) e assicura anche che μ^* realizza il minimo.

Si consideri ora una politica stazionaria μ tale che $P_{\mu^*} J^* = P_{\mu} J^*$; la relazione (2.34) diventa

$$0 \leq g_{\mu} - g_{\mu^*} + \alpha(P_{\mu} - P_{\mu^*})(h^* + O(|1 - \alpha|)).$$

Passando al limite per $\alpha \rightarrow 1$ si ottiene $g_{\mu^*} + P_{\mu^*} h^* \leq g_{\mu} + P_{\mu} h^*$. Per l'equazione (2.23) $J^* + h^* = g_{\mu^*} + P_{\mu^*} h^*$ e dunque $J^* + h^* = g_{\mu^*} + P_{\mu^*} h^* \leq g_{\mu} + P_{\mu} h^*$. Quanto ottenuto è equivalente a (2.32) e dimostra anche che μ^* realizza il minimo. ■

Nel caso particolare in cui il costo medio per slot $J^*(i)$ delle politiche Blackwell ottime sia indipendente dallo stato iniziale i , l'equazione (2.31) risulta automaticamente soddisfatta e si ha $\widehat{\mathcal{A}}(i) = \mathcal{A}(i)$ per ogni i . Tuttavia in generale $\widehat{\mathcal{A}}(i) \neq \mathcal{A}(i)$.

Il seguente teorema consente di semplificare le equazioni (2.31) e (2.32) ed è fondamentale per la dimostrazione dell'ottimalità secondo il criterio del costo medio per slot delle politiche Blackwell ottime.

Teorema 2.4.5 *Si assuma che la coppia (J^*, h^*) soddisfi le equazioni*

$$J^*(i) = \min_{a \in \mathcal{A}(i)} \sum_{j \in \mathcal{S}} p_{ij}(a) J^*(j), \quad i \in \mathcal{S}, \quad (2.35)$$

e

$$J^*(i) + h^*(i) = \min_{a \in \widehat{\mathcal{A}}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h^*(j) \right], \quad i \in \mathcal{S}, \quad (2.36)$$

dove, per ogni i , $\widehat{\mathcal{A}}(i)$ rappresenta l'insieme delle azioni che realizzano il minimo in (2.35). Inoltre, per ogni scalare γ , sia

$$h_\gamma = h^* + \gamma J^*.$$

Esiste allora un qualche $\bar{\gamma} \geq 0$ tale che per ogni $\gamma \geq \bar{\gamma}$ si ha

$$J^*(i) + h_\gamma(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h_\gamma(j) \right], \quad i \in \mathcal{S}. \quad (2.37)$$

Inoltre, se una politica stazionaria μ è tale che $\mu(i)$ realizza il minimo nell'equazione (2.36) per ogni i , allora $\mu(i)$ realizza il minimo in (2.37) per ogni i .

Dimostrazione. L'idea su cui si basa la dimostrazione è quella di aggiungere la penalità

$$\gamma \sum_{j \in \mathcal{S}} p_{ij}(a) J^*(j)$$

alla quantità minimizzata in equazione (2.36) in modo da rendere le azioni $a \notin \widehat{\mathcal{A}}(i)$ poco convenienti nella minimizzazione in (2.37) per un γ sufficientemente grande.

Sia allora μ una politica stazionaria tale che, per ogni i , $\mu(i)$ realizza il minimo nell'equazione (2.36). In base a (2.36) si ha

$$J^*(i) + h^*(i) - \left(g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h^*(j) \right) \leq 0, \quad i \in \mathcal{S}, a \in \widehat{\mathcal{A}}(i), \quad (2.38)$$

dove la disuguaglianza diventa un'uguaglianza quando $a = \mu(i)$.

Utilizzando la definizione $h_\gamma = h^* + \gamma J^*$, si può scrivere

$$\begin{aligned} J^*(i) + h_\gamma(i) - \left(g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h_\gamma(j) \right) &= J^*(i) + h^*(i) \\ &- \left(g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h^*(j) \right) + \gamma \left(J^*(i) - \sum_{j \in \mathcal{S}} p_{ij}(a) J^*(j) \right). \end{aligned} \quad (2.39)$$

Per ogni i e per $a \notin \hat{\mathcal{A}}(i)$ vale

$$J^*(i) - \sum_{j \in \mathcal{S}} p_{ij}(a) J^*(j) < 0,$$

dunque, in base all'equazione (2.39), si può concludere che esiste uno scalare positivo $\gamma(i, a)$ sufficientemente grande tale che

$$J^*(i) + h_\gamma(i) - \left(g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h_\gamma(j) \right) \leq 0, \quad \forall \gamma \geq \gamma(i, a).$$

Per ogni i e per $a \in \hat{\mathcal{A}}(i)$ si ha, invece, $J^*(i) - \sum_{j \in \mathcal{S}} p_{ij}(a) J^*(j) = 0$, quindi utilizzando le equazioni (2.38) e (2.39) si ottiene

$$J^*(i) + h_\gamma(i) - \left(g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h_\gamma(j) \right) \leq 0, \quad \forall \gamma \geq 0,$$

dove la disuguaglianza diventa un'uguaglianza per $a = \mu(i)$. Ma allora fissando $\bar{\gamma} = \max \left\{ \gamma(i, a) \mid i \in \mathcal{S}, a \notin \hat{\mathcal{A}}(i) \right\}$ si ottiene che per ogni $\gamma \geq \bar{\gamma}$ vale

$$J^*(i) + h_\gamma(i) - \left(g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h_\gamma(j) \right) \leq 0, \quad i \in \mathcal{S}, a \in \mathcal{A}(i),$$

dove la disuguaglianza diventa un'uguaglianza per $a = \mu(i)$. Ma questo è proprio il risultato che si voleva dimostrare. ■

Secondo il teorema 2.4.4 la coppia gain-bias (J^*, h^*) comune a tutte le politiche Blackwell ottime soddisfa le equazioni (2.35) e (2.36) e una generica politica Blackwell ottima μ^* ne realizza il minimo. Ma allora per il teorema 2.4.5 è soddisfatta anche l'equazione (2.37) e μ^* realizza il minimo anche in tale equazione.

Nella teoria dei processi decisionali di Markov molto spesso viene introdotta ed utilizzata una funzione mappa, indicata con T , che può essere adottata per ottenere una notazione più compatta [1, 2]. La mappa T è definita nel seguente modo:

$$(Th)(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a)h(j) \right], \quad i \in \mathcal{S}.$$

Considerata poi una politica stazionaria μ si introduce anche la mappa T_μ definita come

$$(T_\mu h)(i) = g(i, \mu(i)) + \sum_{j \in \mathcal{S}} p_{ij}(\mu(i))h(j), \quad i \in \mathcal{S}.$$

Con queste definizioni l'equazione (2.23) diventa ad esempio

$$J_\mu + h_\mu = T_\mu h_\mu.$$

Si definisce inoltre T^N la composizione di T per se stessa N volte e in modo analogo si indica con T_μ^N la composizione di T_μ . Si ha semplicemente:

$$(T^N h)(i) = (T(T^{N-1}h))(i), \quad (T_\mu^N h)(i) = (T_\mu(T_\mu^{N-1}h))(i), \quad i \in \mathcal{S}.$$

È facile dimostrare per induzione [2] che $(T^N h)(i)$ è il costo totale ottimo per un problema decisionale di Markov ad orizzonte finito N con stato iniziale i , costi g sul singolo slot e funzione costo finale h , dove per costo finale si intende il costo accumulato nell'istante finale N in base allo stato occupato. Allo stesso modo $(T_\mu^N h)(i)$ è il costo totale della politica stazionaria μ per lo stesso problema. Formalmente:

$$(T_\mu^N h)(i) = E \left[h(s_N) + \sum_{t=0}^{N-1} g(s_t, \mu(s_t)) \middle| s_0 = i \right], \quad i \in \mathcal{S}. \quad (2.40)$$

Le definizioni appena introdotte possono essere facilmente estese nel caso in cui si consideri una generica politica ammissibile $\pi = \{\mu_0, \mu_1, \dots, \mu_{N-1}\}$. L'espressione $(T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}} h)(i)$ può essere definita in modo ricorsivo tenendo conto che per ogni $x = 0, 1, \dots, N-2$ si ha

$$(T_{\mu_x} T_{\mu_{x+1}} \cdots T_{\mu_{N-1}} h)(i) = (T_{\mu_x} (T_{\mu_{x+1}} \cdots T_{\mu_{N-1}} h))(i).$$

Anche in questo caso $(T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}} h)(i)$ rappresenta il costo totale della politica π per un problema decisionale ad orizzonte finito N con stato iniziale i , costi g sul singolo slot e funzione costo finale h .

Una caratteristica fondamentale delle mappe T^N e T_μ^N è quella di essere *monotone* per ogni μ e $N \geq 1$.

Quanto detto consente di introdurre il seguente teorema.

Teorema 2.4.6 *Siano J e h due vettori di dimensione S e sia $\pi = \{\mu_0, \mu_1, \dots\}$ una generica politica ammissibile. Se per ogni $t = 0, 1, 2, \dots$ valgono*

$$P_{\mu_t} J \geq J, \quad T_{\mu_t} h \geq J + h, \quad (2.41)$$

allora

$$J_\pi \geq J. \quad (2.42)$$

Inoltre, se le (2.41) valgono come uguaglianze per ogni t , allora $J_\pi = J$.

Dimostrazione. Per ogni μ si ha $T_\mu(J + h) = g_\mu + P_\mu(J + h) = P_\mu J + T_\mu h$. In base alle relazioni (2.41) segue allora che per ogni $t = 0, 1, 2, \dots$ vale

$$T_{\mu_t}(J + h) \geq J + T_{\mu_t} h. \quad (2.43)$$

Sia ora N un intero positivo. Sempre per le (2.41) si ha

$$T_{\mu_{N-1}} h \geq J + h. \quad (2.44)$$

Applicando $T_{\mu_{N-2}}$ ad entrambi i membri di (2.44), sfruttando la monotonicità di $T_{\mu_{N-2}}$ e utilizzando le relazioni (2.41) e (2.43) si ottiene

$$T_{\mu_{N-2}} T_{\mu_{N-1}} h \geq T_{\mu_{N-2}}(J + h) \geq J + T_{\mu_{N-2}} h \geq 2J + h.$$

Se ora si applica $T_{\mu_{N-3}}$ e si procede allo stesso modo fino ad applicare T_{μ_0} si ricava la disuguaglianza

$$T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}} h \geq NJ + h, \quad (2.45)$$

che vale come uguaglianza se le (2.41) sono soddisfatte con uguaglianza per ogni t .

Per ogni $i \in \mathcal{S}$, $(T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}} h)(i)$ rappresenta il costo totale della politica $\{\mu_0, \mu_1, \dots, \mu_{N-1}\}$ per un problema decisionale ad orizzonte finito N con stato iniziale i e funzione costo finale h . Pertanto, analogamente a (2.40), si può scrivere

$$(T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}} h)(i) = E \left[h(s_N) + \sum_{t=0}^{N-1} g(s_t, \mu_t(s_t)) \middle| s_0 = i, \pi \right], \quad i \in \mathcal{S}.$$

Utilizzando questa espressione in (2.45) e dividendo per N si ottiene

$$\frac{1}{N}E[h(s_N)|s_0 = i, \pi] + \frac{1}{N}E\left[\sum_{t=0}^{N-1} g(s_t, \mu_t(s_t))\middle|s_0 = i, \pi\right] \geq J(i) + \frac{1}{N}h(i), \quad i \in \mathcal{S}.$$

Eseguendo il limite superiore per $N \rightarrow \infty$ si evince che

$$J_\pi(i) \geq J(i), \quad i \in \mathcal{S}.$$

Dunque la relazione (2.42) è dimostrata.

Infine se le (2.41) sono soddisfatte con uguaglianza per ogni t , tutte le disuguaglianze incontrate nel procedimento appena concluso valgono come uguaglianze. Ripetendo il ragionamento si giunge quindi a dimostrare che $J_\pi = J$. ■

Il teorema appena dimostrato fornisce una condizione sufficiente per l'ottimalità di una politica stazionaria secondo il criterio del costo medio per slot. Infatti, se μ è una politica stazionaria ed esistono due vettori J e h tali per cui valgono

$$J = P_\mu J = \min_{\mu'} P_{\mu'} J, \quad J + h = T_\mu h = \min_{\mu'} T_{\mu'} h, \quad (2.46)$$

allora, per il teorema 2.4.6, μ è ottima e J è il costo medio per slot ottimo. Questa osservazione è fondamentale per la dimostrazione della seguente proposizione.

Proposizione 2.4.7 *Una politica Blackwell ottima è ottima secondo il criterio del costo medio per slot tra tutte le politiche ammissibili.*

Dimostrazione. Sia μ una politica stazionaria Blackwell ottima con coppia gain-bias (J_μ, h_μ) . Si ponga

$$J = J_\mu, \quad h = h_\mu + \gamma J,$$

dove γ è uno scalare sufficientemente grande affinché valga $Th = J + h$, in base al teorema 2.4.5. Allora, per qualsiasi politica stazionaria μ' , si ha

$$T_{\mu'} h \geq J + h.$$

D'altra parte $J = J_\mu$ soddisfa l'equazione (2.35) e dunque si ha anche $P_{\mu'} J \geq J$ per ogni μ' . Ma questo implica che per qualsiasi politica ammissibile $\pi = \{\mu_0, \mu_1, \dots\}$

le disuguaglianze (2.41) sono soddisfatte per ogni t . Il teorema 2.4.6 assicura quindi che $J_\pi \geq J = J_\mu$ per ogni politica π . Pertanto, la politica μ è ottima. ■

Dalla proposizione 2.4.7 discende direttamente la seguente.

Proposizione 2.4.8 *Esiste una politica stazionaria ottima secondo il criterio del costo medio per slot per ogni problema decisionale di Markov.*

Dimostrazione. In base alla proposizione 2.4.3, per ogni problema decisionale di Markov esiste una politica stazionaria Blackwell ottima. Ma, per la proposizione 2.4.7, una politica Blackwell ottima è ottima secondo il criterio del costo medio per slot. ■

2.4.4 Equazioni di ottimalità

I risultati fin qui ottenuti consentono di introdurre le equazioni di ottimalità valide nel caso si consideri come criterio di ottimizzazione il costo medio per slot.

Definizione 2.4.2 *Si definiscono equazioni di ottimalità, note anche come equazioni di Bellman, per il costo medio per slot*

$$J(i) = \min_{a \in \mathcal{A}(i)} \sum_{j \in \mathcal{S}} p_{ij}(a) J(j), \quad i \in \mathcal{S}, \quad (2.47)$$

e

$$J(i) + h(i) = \min_{a \in \hat{\mathcal{A}}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h(j) \right], \quad i \in \mathcal{S}, \quad (2.48)$$

dove, per ogni i , $\hat{\mathcal{A}}(i)$ rappresenta l'insieme delle azioni che realizzano il minimo in (2.47).

Per il teorema 2.4.4 la coppia gain-bias (J^*, h^*) comune a tutte le politiche Blackwell ottime soddisfa le equazioni di ottimalità. Inoltre, la proposizione 2.4.3 assicura che esiste sempre una politica Blackwell ottima. Ma allora le equazioni di Bellman possiedono sempre almeno una soluzione.

Il risultato che porta a definire le relazioni (2.47) e (2.48) equazioni di ottimalità è il seguente.

Proposizione 2.4.9 *Se i vettori J^* e h^* soddisfano le equazioni di ottimalità (2.47) e (2.48), allora J^* è il vettore ottimo dei costi medi per slot. Inoltre, se $\mu^*(i)$ realizza il minimo nelle equazioni (2.47) e (2.48) per ogni $i \in \mathcal{S}$, allora la politica stazionaria μ^* è ottima secondo il criterio del costo medio per slot.*

Dimostrazione. Dato che J^* e h^* soddisfano (2.47) e (2.48), in base al teorema 2.4.5 è possibile fissare uno scalare γ in modo che, posto $h_\gamma = h^* + \gamma J^*$, sia soddisfatta

$$J^*(i) + h_\gamma(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h_\gamma(j) \right], \quad i \in \mathcal{S}. \quad (2.49)$$

Si assuma ora che $\mu^*(i)$ realizzi il minimo nelle equazioni (2.47) e (2.48) per ogni i . Sempre per il teorema 2.4.5, $\mu^*(i)$ minimizza anche l'equazione (2.49) per ogni i .

Utilizzando il teorema 2.4.6 con $J = J^*$ e $h = h_\gamma$ si ha allora che per ogni politica ammissibile $\pi = \{\mu_0, \mu_1, \dots\}$ le relazioni (2.41) sono soddisfatte per ogni t . Questo consente di concludere che $J_\pi \geq J = J^*$ per ogni politica π . Applicando nuovamente il teorema 2.4.6 con $J = J^*$, $h = h_\gamma$ e politica $\pi = \{\mu^*, \mu^*, \dots\}$, le relazioni (2.41) valgono come uguaglianze per ogni t ; dunque $J_{\mu^*} = J = J^*$.

Complessivamente si ottiene

$$J_\pi \geq J^* = J_{\mu^*},$$

per ogni politica ammissibile π . Quindi J^* è il vettore ottimo e μ^* è ottima secondo il criterio del costo medio per slot. ■

Un caso particolare molto importante si ha quando il costo medio per slot ottimo è il medesimo per qualsiasi stato iniziale; si dice, in questo caso, che il costo medio per slot ottimo è *costante*. Esiste allora un qualche scalare λ tale che

$$J(i) = \lambda, \quad \forall i \in \mathcal{S}. \quad (2.50)$$

Se è valida la condizione (2.50), l'equazione (2.47) è implicitamente soddisfatta e diventa superflua; si ha, quindi, $\mathcal{A}(i) = \widehat{\mathcal{A}}(i)$ per ogni i . Ma allora le equazioni di ottimalità (2.47) e (2.48) si riducono alla singola equazione

$$\lambda + h(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h(j) \right], \quad i \in \mathcal{S}. \quad (2.51)$$

Anche l'equazione (2.51) è detta equazione di ottimalità o equazione di Bellman per il costo medio per slot, sottintendendo che essa è valida solo quando il costo medio per slot ottimo è costante.

A questo punto si dimostra facilmente il seguente risultato.

Proposizione 2.4.10 *Se esistono uno scalare λ e un vettore h che soddisfano*

$$\lambda + h(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h(j) \right], \quad i \in \mathcal{S}, \quad (2.52)$$

allora λ è il costo medio per slot ottimo per ogni i , cioè

$$\lambda = \min_{\pi \in \Pi} J_{\pi}(i) = J(i), \quad i \in \mathcal{S}.$$

Inoltre, se $\mu^(i)$ realizza il minimo nell'equazione (2.52) per ogni i , allora la politica stazionaria μ^* è ottima e $J_{\mu^*}(i) = \lambda$ per ogni i .*

Dimostrazione. Segue direttamente come caso particolare della proposizione 2.4.9. ■

Infine, è immediato verificare che vale anche la seguente.

Proposizione 2.4.11 *Si consideri una politica stazionaria μ . Se esistono uno scalare λ_{μ} e un vettore h tali che*

$$\lambda_{\mu} + h(i) = g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(\mu(i)) h(j), \quad i \in \mathcal{S},$$

allora $\lambda_{\mu} = J_{\mu}(i)$ per ogni i .

3

Modello del sensore radio

Gli strumenti matematici offerti dalla teoria dei processi decisionali di Markov sono ampiamente sfruttati in questo capitolo dedicato a quello che è il principale contributo della tesi: il modello matematico di un sensore radio alimentato da una fonte di energia rinnovabile. Prima di tutto si individuano le componenti principali che contraddistinguono un sistema di questo tipo e si fornisce una rappresentazione formale per ciascuna di esse. Combinando insieme i modelli dei singoli elementi costitutivi del dispositivo si definisce poi il modello complessivo. In definitiva, quello che si propone è un processo decisionale di Markov a tempo discreto tale da poter essere trattato dal punto di vista computazionale nella fase di ottimizzazione successiva.

3.1 L'idea

Il sistema preso in considerazione è un generico sensore radio alimentato da una qualche fonte di energia rinnovabile, che collabora con altri nodi all'interno di una WSN. Per realizzare un modello in grado di rappresentare un dispositivo di questo tipo, piuttosto complicato, l'idea è quella di individuare gli aspetti fondamentali che vale la pena considerare. Vi è la necessità di contenere il più possibile la complessità del modello, consapevoli del fatto che su di esso dovrà essere applicata una procedura di ottimizzazione con i relativi problemi computazionali. Al contempo però vi è anche l'esigenza di catturare le dinamiche caratteristiche del sistema analizzato.

Alla luce di questo si schematizza il dispositivo e il suo funzionamento come mostrato in figura 3.1. In questo diagramma a blocchi sono rappresentate le principali grandezze in gioco, gli elementi costitutivi del sensore radio e il legame che intercorre tra essi.

Le componenti principali, comprese nello schema, su cui si concentra l'attenzione sono le seguenti:

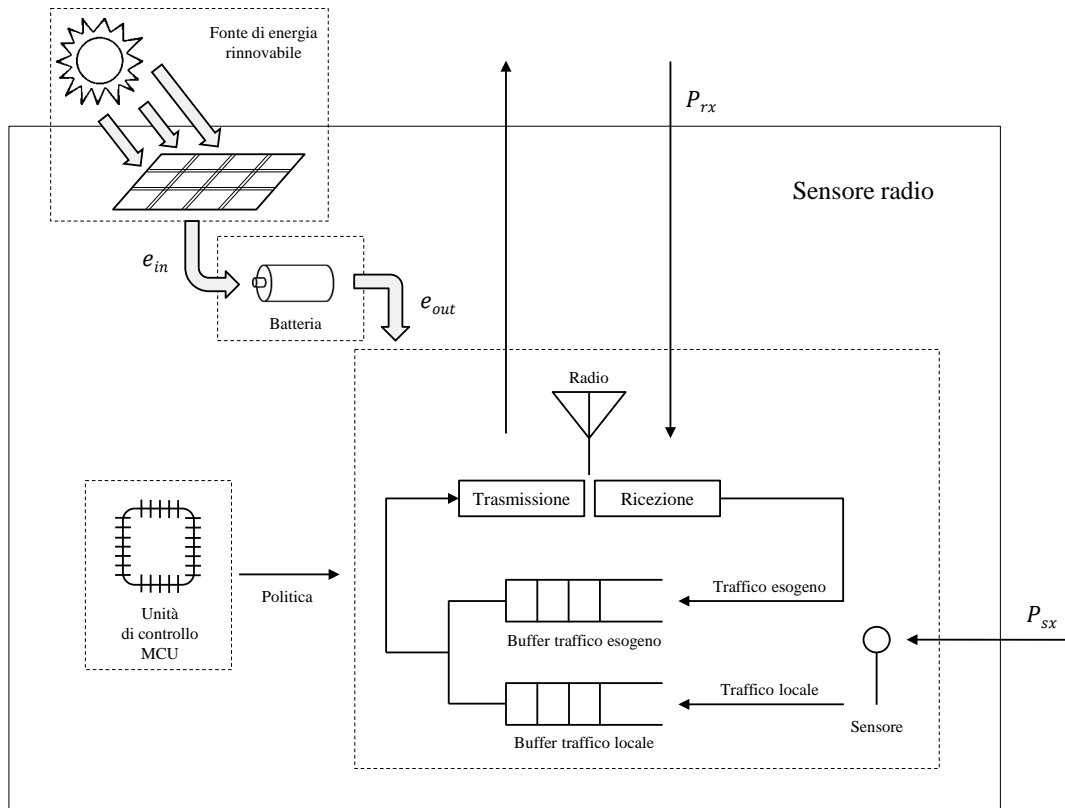


Figura 3.1 Schema del sensore radio.

- **Batteria:** fornisce l'energia necessaria per le operazioni compiute dal dispositivo. Col termine batteria si indica in generale un qualsiasi componente in grado di immagazzinare localmente l'energia. Nelle implementazioni reali è possibile utilizzare ad esempio batterie ricaricabili nichel metallo idruro (NiMH) oppure agli ioni di litio (Li-ion) o, ancora, batterie nichel cadmio (NiCd); una possibile alternativa è anche l'adozione di supercapacitori o supercondensatori. Ciò che è richiesto alla batteria è esclusivamente la capacità di alimentare il sensore radio e la possibilità di ricaricarsi mediante l'energia acquisita dall'ambiente circostante.
- **Radio:** consente la ricezione e la trasmissione delle informazioni da e verso gli altri nodi della WSN. La radio può essere spenta o accesa; quando è attiva consente la ricezione di pacchetti dall'esterno *oppure* la trasmissione di quelli contenuti nei buffer. Le due azioni, quindi, sono mutuamente esclusive.
- **Sensore o unità di sensing:** si occupa della rilevazione e della misurazione dei

dati ambientali di interesse per l'applicazione.

- *Buffer per il traffico locale*: è il supporto di memorizzazione per i dati provenienti dal sensore e che quindi sono generati localmente al dispositivo. Tali informazioni vengono salvate in questo buffer in attesa di poter essere trasmesse dalla radio.
- *Buffer per il traffico esogeno*: memorizza le informazioni provenienti dagli altri nodi della WSN fino a quando possono essere ritrasmesse dalla radio.
- *Unità di controllo o MCU (MicroController Unit)*: gestisce l'attività svolta dal sensore radio specificando di volta in volta quale azione debba essere eseguita secondo una determinata *politica di scheduling*. Una politica di scheduling stabilisce qual è l'operazione che deve essere compiuta dato che il dispositivo si trova in una determinata condizione; è esattamente lo stesso concetto di politica introdotto in modo formale nel capitolo 2.
- *Fonte di energia rinnovabile*: è la sorgente che fornisce l'energia mediante la quale la batteria si ricarica. La figura 3.1 si riferisce per chiarezza all'energia solare ma si può considerare, in generale, una qualsiasi altra fonte di energia rinnovabile.

Le tipiche operazioni che un sensore radio di una WSN deve compiere sono:

- *Rilevazione o sensing*: il dispositivo, attraverso il sensore, monitora il fenomeno fisico d'interesse nell'ambiente circostante. Una volta eseguita la rilevazione, se le informazioni acquisite sono nuove rispetto a quelle ottenute dall'attività di sensing precedente allora vengono memorizzate nel buffer per il traffico locale, altrimenti vengono ignorate.
- *Inoltro o relay*: il sensore radio è responsabile di inoltrare le informazioni ricevute dagli altri nodi della rete in modo che queste possano raggiungere il destinatario voluto. Ad esempio potrebbe essere necessario ritrasmettere dei dati provenienti da un altro dispositivo affinché possano essere ricevuti dal *sink*.¹ L'operazione

¹Il sink è un nodo speciale della rete che si preoccupa di raccogliere le informazioni sulle rilevazioni compiute e di comunicarle all'esterno della WSN.

di inoltre prevede innanzitutto di utilizzare la radio per ricevere le informazioni; quest'ultime poi vengono memorizzate nel buffer per il traffico esogeno in attesa di essere ritrasmesse.

Durante la sua evoluzione nel tempo, il sensore radio svolge le azioni indicate dall'unità di controllo in base ad una certa politica di scheduling. Ogni azione compiuta porta al consumo di una determinata quantità di energia che dipende da quali componenti vengono attivati; naturalmente tale energia viene prelevata dalla batteria. Parallelamente la batteria può però contare sull'apporto energetico fornito dalla fonte rinnovabile.

3.2 Assunzioni

Il diagramma a blocchi introdotto in figura 3.1 è il punto di partenza per la costruzione del modello. Quest'ultimo viene realizzato in base anche ad alcune assunzioni che si supporranno soddisfatte nel prosieguo e che vengono qui evidenziate:

- *Tempo suddiviso in slot*: il tempo è suddiviso e misurato in periodi o slot di durata prefissata e costante nel tempo. Si assume quindi di rappresentare il sensore radio come un sistema a tempo discreto.
- *Pacchetti dati*: le informazioni ricevute dagli altri nodi, quelle trasmesse, quelle originate dall'attività di sensing e quelle memorizzate nei buffer sono gestite dal dispositivo sotto forma di pacchetti dati di dimensione fissata e determinata in base alla particolare applicazione. Se in un determinato slot il dispositivo attiva la radio in trasmissione, le informazioni inviate agli altri nodi sono contenute in un pacchetto; allo stesso modo i dati ottenuti da un'operazione di ricezione o rilevazione compiuta in uno slot formano un pacchetto che viene memorizzato nel buffer opportuno.
- *Azioni completate nello slot*: le azioni che il sensore radio sceglie di svolgere occupano l'intera durata di uno slot e vengono eseguite completamente in tale intervallo temporale. Ciò significa che la durata dello slot deve essere tale da consentire la ricezione o il sensing o la trasmissione di un intero pacchetto dati.

- *Unità di carica*: l'energia richiesta e fornita alla batteria e quella in essa contenuta è espressa in termini di unità di carica; un'unità di carica è la minima quantità di energia trasferibile. Ogni operazione eseguita dal dispositivo necessita di un certo numero di unità di carica che vengono prelevate dalla batteria.
- *Trasferimenti di energia effettivi*: quando il dispositivo compie un'azione, il numero di unità di carica sottratte dalla batteria comprende i consumi di tutti i componenti circuitali necessari per portare a termine l'operazione. Allo stesso modo le unità di carica fornite alla batteria dalla fonte rinnovabile sono considerate al netto di tutte le inefficienze di conversione.
- *Stazionarietà*: le caratteristiche strutturali del dispositivo rimangono invariate nel tempo.

3.3 Obiettivi

Alla luce delle idee e delle assunzioni viste nelle sezioni 3.1 e 3.2, il primo obiettivo è ottenere dallo schema logico generale di figura 3.1 un modello stocastico per il sensore radio. L'intenzione è quella di rappresentare il dispositivo mediante un processo decisionale di Markov a tempo discreto che soddisfi le ipotesi presentate in sezione 2.3.1 e per il quale valgono, dunque, i risultati introdotti in sezione 2.4.

Ottenuto un processo decisionale di Markov rappresentativo del sistema, interessa poi determinare, al variare dei parametri, le politiche (di scheduling) ottime che, una volta applicate dall'unità di controllo, portino il sensore radio a massimizzare le sue prestazioni. Nello specifico si vogliono individuare le politiche che rendano massimo il *throughput*, ovvero il numero medio di pacchetti trasmessi per slot.

Si desidera, inoltre, che il dispositivo possa massimizzare il throughput dando precedenza ad una delle due attività tra quella di rilevazione dei dati ambientali e quella di inoltrare dei pacchetti provenienti da altri nodi. Nel calcolo delle politiche ottime, quindi, deve essere possibile specificare, mediante un opportuno parametro, quale attività ha maggior rilevanza e in che misura. In questo modo è possibile determinare il comportamento ottimo per il sensore radio in base al ruolo che ricopre nella WSN.

Alcuni nodi, infatti, hanno principalmente il compito di monitorare il fenomeno di interesse, altri quello di inoltrare le informazioni nella rete: per i primi avrà precedenza l'attività di sensing, per i secondi quella di relay. Per il calcolo delle politiche che massimizzino il throughput sarà adottato il criterio di ottimizzazione del costo medio per slot di cui si è discusso in sezione 2.4; preliminarmente però dovranno essere definiti in modo opportuno i costi accumulati in ogni singolo slot.

Con questi propositi si considerano a questo punto le componenti principali del sensore radio messe in evidenza nello schema di figura 3.1 e si propone un modello per ciascuna di esse; dalla combinazione di questi modelli specifici si ricaverà quello complessivo.

3.4 Batteria

Il ruolo assunto dalla batteria nel dispositivo è fondamentale in quanto è sua responsabilità fornire l'energia necessaria per compiere le azioni e immagazzinare quella proveniente dalla fonte rinnovabile.

Si propone un modello piuttosto immediato e generale che rappresenta la batteria come un semplice *serbatoio di unità di carica* soggetto ad un flusso di carica entrante e ad uno uscente. La figura 3.2 chiarisce la situazione.

Si indica con N la capacità della batteria intesa come il numero di unità che essa contiene quando è completamente carica. Ciò significa che N sono le unità di carica che la batteria è in grado di fornire prima di esaurirsi completamente nel caso in cui non venga ricaricata.

Lo stato della batteria all'inizio di un generico slot temporale è rappresentato mediante la variabile n che specifica il numero di unità di carica rimaste; naturalmente si ha $n \in \{0, 1, \dots, N\}$. La batteria è esausta quando $n = 0$.

In ogni slot alla batteria viene richiesto un certo numero di unità di carica, indicato con e_{out} , a seconda della particolare azione che il dispositivo decide di eseguire; in base ai componenti attivati per svolgere l'operazione vengono prelevate dalla batteria più o meno unità. Contestualmente la fonte di energia rinnovabile ricarica la batteria fornendole un certo numero e_{in} di unità.

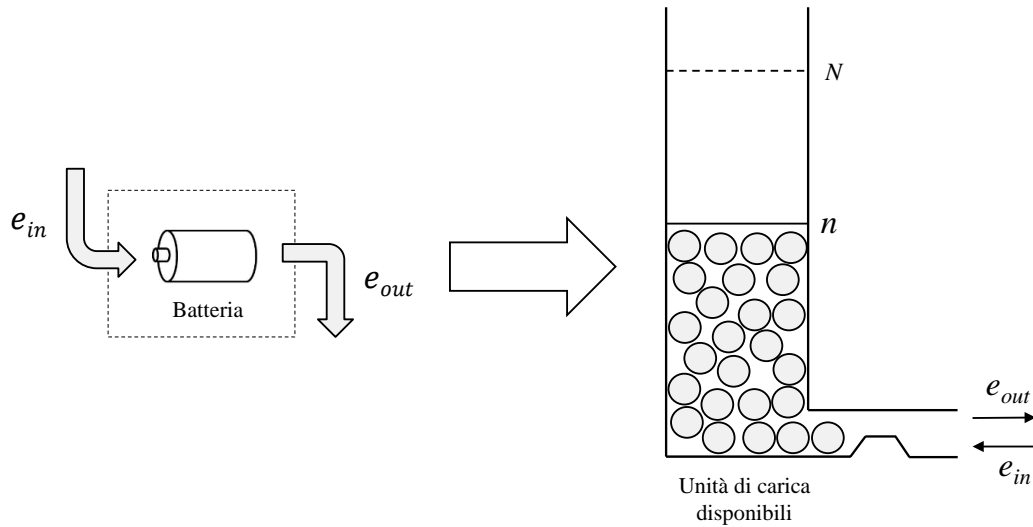


Figura 3.2 Modello a serbatoio della batteria.

Quindi, se lo stato della batteria all'inizio di un generico slot è n , lo stato occupato all'inizio dello slot successivo è dato dalla relazione di aggiornamento

$$n \rightarrow n + e_{in} - e_{out}. \quad (3.1)$$

Si osservi che mentre il valore di e_{out} è determinato dalla particolare azione eseguita dal dispositivo nello slot considerato, quello di e_{in} è aleatorio e condizionato dallo stato in cui si trova la fonte di energia rinnovabile. Nel caso in cui $n + e_{in} - e_{out} > N$, lo stato della batteria nello slot successivo è N e il contributo energetico in eccesso dato dalla fonte rinnovabile viene perso.

3.4.1 Livello di guardia

Può essere utile evitare che la batteria si scarichi oltre una certa soglia; infatti mantenere il livello di carica abbastanza elevato consente di prolungare il tempo di vita della batteria. Tipicamente interessa mantenere il numero di unità di carica residue sopra una qualche frazione ρ della capacità totale. Si fissa quindi il *livello di guardia*

$$N_{lim} = \lceil \rho N \rceil, \quad (3.2)$$

con $0 \leq \rho \leq 1$. Allora, durante l'evoluzione del sistema, si vorrebbe fosse sempre soddisfatta la condizione $n \geq N_{lim}$.

3.5 Fonte di energia rinnovabile

Nello schema di riferimento di figura 3.1 si indica con fonte di energia rinnovabile l'insieme composto sia dalla sorgente di energia vera e propria (sole, vento, ...) che dai componenti responsabili della conversione dell'energia ambientale in energia elettrica che viene poi utilizzata per ricaricare la batteria. Alla luce di questo, si suppone che in ciascuno slot temporale la fonte di energia rinnovabile possa fornire direttamente alla batteria un certo numero di unità di carica indicato con e_{in} . A questo punto vi è quindi la necessità di introdurre un modello che fornisca una caratterizzazione statistica di questo apporto energetico su cui la batteria può contare. Si introduce prima di tutto un modello di carattere generale che possa essere applicato qualsiasi sia la sorgente di energia ambientale; successivamente lo si particularizza al caso in cui si utilizzi l'energia solare, essendo molto comune nelle implementazioni reali.

3.5.1 Modello generale

L'idea di base è quella di rappresentare la fonte rinnovabile mediante una *catena di Markov a stati finiti* in cui i vari stati indichino diverse intensità del flusso energetico che ricarica la batteria. Si faccia riferimento alla figura 3.3.

In un qualsiasi slot temporale la fonte fornisce un certo numero e_{in} di unità di carica alla batteria; sia E_{MAX} il massimo numero di unità trasferibili per slot. Il valore di E_{MAX} dipende strettamente dalla particolare sorgente ambientale considerata e dal meccanismo di conversione adottato; ad esempio nel caso dell'energia solare sono rilevanti fattori come la dimensione del pannello solare e l'efficienza di conversione. Naturalmente si ha $e_{in} \in \{0, 1, \dots, E_{MAX}\}$.

Si partiziona ora l'insieme $\{0, 1, \dots, E_{MAX}\}$ in $X + 1$ sottoinsiemi ΔE_x , $x = 0, 1, \dots, X$; come in figura 3.3 si supponga ad esempio $E_{MAX} = 7$ e $X = 3$, allora una possibile suddivisione dell'insieme $\{0, 1, \dots, 7\}$ è $\Delta E_0 = \{0, 1\}$, $\Delta E_1 = \{2, 3\}$, $\Delta E_2 = \{4, 5\}$, $\Delta E_3 = \{6, 7\}$. A ciascun sottoinsieme viene quindi associato uno

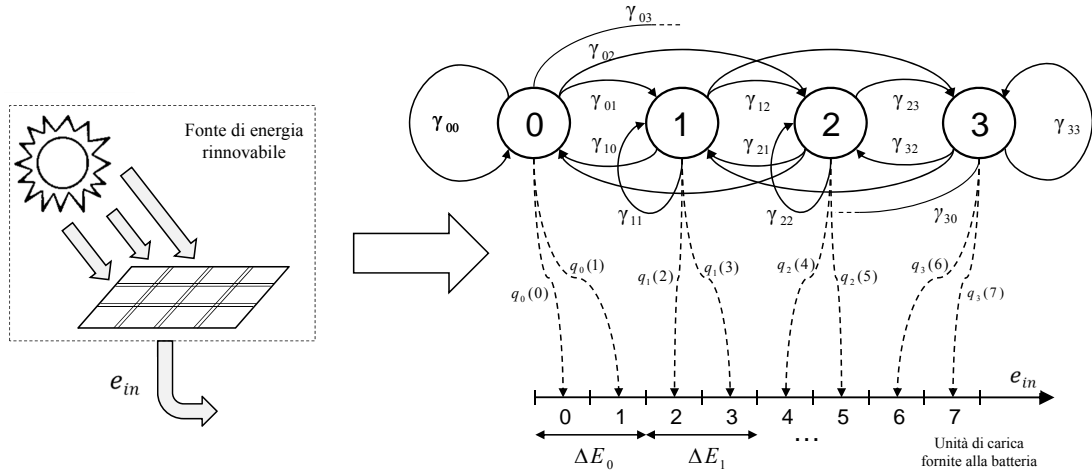


Figura 3.3 Modello generale della fonte di energia rinnovabile con $E_{MAX} = 7$, $X = 3$, $\Delta E_0 = \{0, 1\}$, $\Delta E_1 = \{2, 3\}$, $\Delta E_2 = \{4, 5\}$ e $\Delta E_3 = \{6, 7\}$

stato di una catena di Markov con $X + 1$ stati complessivi; lo stato $x \in \{0, 1, \dots, X\}$ rappresenta ΔE_x . Gli stati della catena di Markov indicano diversi livelli di intensità del flusso energetico dalla fonte alla batteria. Passando dallo stato 0 allo stato 1 fino a giungere allo stato X l'apporto energetico aumenta.

La catena di Markov così ottenuta viene adottata per modellare l'evoluzione della sorgente di energia e il contributo energetico fornito alla batteria. Se in un generico slot lo stato occupato dalla catena è $x \in \{0, 1, \dots, X\}$, si assume che la fonte rinnovabile alimenti con e_{in} unità di carica la batteria secondo la distribuzione $q_x(e_{in})$. Dunque $q_x(e_{in})$ è la probabilità che nello slot corrente la sorgente fornisca e_{in} unità di carica dato che lo stato in cui si trova è x . Le distribuzioni di probabilità $q_x(\cdot)$ soddisfano le condizioni

$$q_x(e) = 0 \text{ per } e \notin \Delta E_x, \quad x = 0, 1, \dots, X, \quad (3.3)$$

e

$$\sum_{e \in \Delta E_x} q_x(e) = 1, \quad q_x(e) \in [0, 1] \quad \forall e, \quad x = 0, 1, \dots, X. \quad (3.4)$$

La condizione (3.3) assicura che se la fonte di energia si trova nello stato x il numero di unità di carica fornite alla batteria appartenga all'intervallo ΔE_x rappresentato da x ; con la (3.4), invece, si impone che le funzioni $q_x(\cdot)$ soddisfino la condizione di normalizzazione e che quindi siano distribuzioni di probabilità ben definite.

Le probabilità di transizione tra gli $X + 1$ stati della catena sono invece specificate dalla matrice

$$\Gamma = \begin{bmatrix} \gamma_{00} & \gamma_{01} & \gamma_{02} & \gamma_{03} & \cdots & \gamma_{0X} \\ \gamma_{10} & \gamma_{11} & \gamma_{12} & \cdots & & \gamma_{1X} \\ \gamma_{20} & \gamma_{21} & \ddots & & & \\ \vdots & \vdots & & & & \\ & & & & \ddots & \\ \vdots & \vdots & & & & \\ \gamma_{X0} & \gamma_{X1} & \cdots & & & \gamma_{XX} \end{bmatrix}.$$

Tali probabilità possono essere calcolate empiricamente a partire da profili energetici reali della fonte rinnovabile utilizzata. Si consideri per semplicità la radiazione solare e si supponga di misurare il contributo energetico fornito alla batteria nel corso di un determinato periodo di tempo; nel caso specifico può essere sensato considerare ad esempio un'intera giornata. Per determinare la probabilità di transizione $\gamma_{xx'}$ dallo stato x a x' si può calcolare il rapporto tra il numero di transizioni nel profilo energetico che avvengono dall'intervallo ΔE_x a $\Delta E_{x'}$, tra uno slot e quello successivo, e il numero di transizioni totali da ΔE_x verso un qualsiasi altro intervallo.

Nel complesso, quindi, la dinamica della fonte di energia rinnovabile è così modellata: se x è lo stato occupato in un generico slot temporale, in quello successivo la sorgente di energia si troverà nello stato x' con probabilità $\gamma_{xx'}$ e il numero di unità di carica fornite alla batteria sarà e_{in} con probabilità $q_{x'}(e_{in})$. Si noti che il modello può essere utilizzato per qualsiasi fonte rinnovabile impostando opportunamente E_{MAX} , le distribuzioni $q_x(\cdot)$ e le probabilità di transizione in Γ .

3.5.2 Modello generale a stato singolo

Il modello generale proposto in sezione 3.5.1 contiene come caso particolare un interessante *modello a singolo stato* ottenibile ponendo $X = 0$. Come mostrato in figura 3.4, la catena di Markov rappresentativa della fonte di energia rinnovabile si riduce al solo stato 0 e, conseguentemente, $\Delta E_0 = \{0, 1, \dots, E_{MAX}\}$. Adottare un

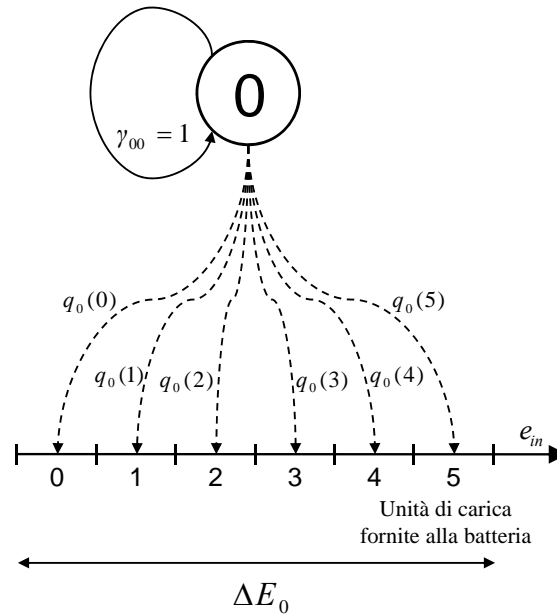


Figura 3.4 Modello a stato singolo ($X = 0$) con $E_{MAX} = 5$ e $\Delta E_0 = \{0, 1, 2, 3, 4, 5\}$

modello di questo tipo significa assumere che in ciascuno slot la batteria riceva dalla sorgente un numero di unità di carica determinato dalla distribuzione $q_0(\cdot)$. In ogni slot, cioè, vengono fornite alla batteria $e_{in} \in \Delta E_0 = \{0, 1, \dots, E_{MAX}\}$ unità di carica con probabilità $q_0(e_{in})$. In questo caso quindi la sorgente di energia è completamente caratterizzata dalla sola distribuzione $q_0(\cdot)$.

3.5.3 Energia solare: modello a due stati

Se la fonte di energia rinnovabile considerata è la radiazione solare il modello generale di sezione 3.5.1 può essere particolarizzato tenendo conto dell'intrinseca periodicità di tale sorgente: durante le ore notturne la radiazione è assente e il contributo energetico è nullo o notevolmente ridotto, mentre nelle ore diurne la radiazione è presente con intensità variabile a seconda del momento considerato.

In questo caso una prima possibilità è quella di adottare un *modello a due stati*, rappresentato in figura 3.5, ricavato dal modello generale fissando $X = 1$ e ponendo $\Delta E_0 = \{0\}$ e $\Delta E_1 = \{1, 2, \dots, E_{MAX}\}$. Allora se la catena occupa lo stato 0 il numero di unità di carica fornite alla batteria è nullo; lo stato 0, cioè, rappresenta l'assenza di radiazione solare nelle ore notturne e la conseguente incapacità della sorgente di fornire

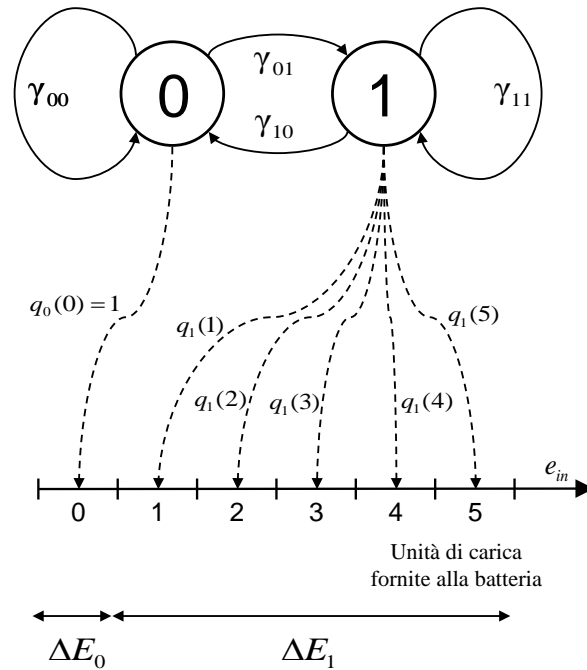


Figura 3.5 Modello per l'energia solare a due stati ($X = 1$) con $E_{MAX} = 5$, $\Delta E_0 = \{0\}$ e $\Delta E_1 = \{1, 2, 3, 4, 5\}$.

energia alla batteria. Lo stato 1 indica, invece, che è presente radiazione solare e il numero di unità di carica acquisite dalla batteria è in accordo alla distribuzione $q_1(\cdot)$ opportunamente definita.

Conseguentemente, se in un generico slot la fonte di energia si trova nello stato 0 la batteria acquisisce $e_{in} = 0$ unità di carica con probabilità $q_0(0) = 1$; se, invece, occupa lo stato 1 vengono fornite $e_{in} \in \Delta E_1 = \{1, 2, \dots, E_{MAX}\}$ unità con probabilità $q_1(e_{in})$.

Una possibile alternativa è adottare il modello a due stati più generale illustrato in figura 3.6. In questo caso si assume $\Delta E_0 = \Delta E_1 = \{0, 1, 2, \dots, E_{MAX}\}$. Ciò che cambia sono le distribuzioni $q_0(\cdot)$ e $q_1(\cdot)$ relative ai due stati: è opportuno che la distribuzione $q_0(\cdot)$ sia molto sbilanciata verso valori bassi dell'insieme $\{0, 1, 2, \dots, E_{MAX}\}$, mentre la distribuzione $q_1(\cdot)$ può essere definita in modo da favorire soprattutto i valori maggiori di 0. In questo modo, quando la fonte rinnovabile si trova nello stato 0, che rappresenta le ore notturne, la batteria viene ricaricata solo con un numero esiguo di unità, mentre se lo stato occupato è 1 l'apporto energetico risulta più significativo.

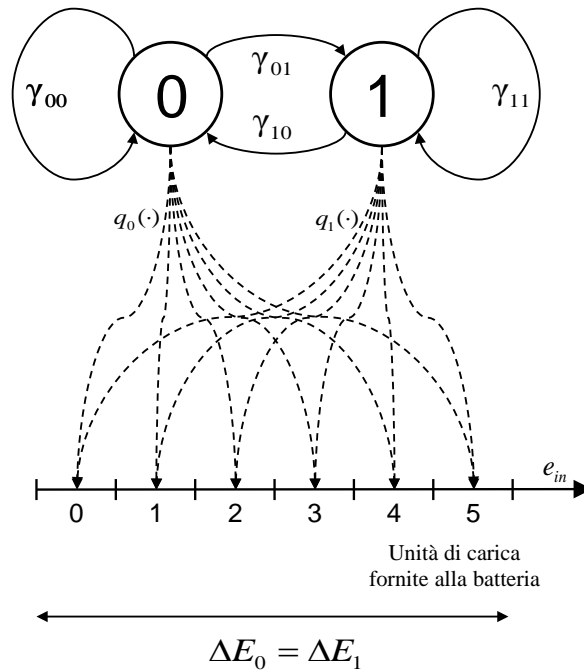


Figura 3.6 Modello a due stati alternativo per l'energia solare con $E_{MAX} = 5$, $\Delta E_0 = \Delta E_1 = \{0, 1, 2, 3, 4, 5\}$ e distribuzioni $q_0(\cdot)$ e $q_1(\cdot)$ distinte.

3.5.4 Energia solare: modello a più stati

Il modello di sezione 3.5.3 presentato in figura 3.5 si può estendere introducendo più stati ($X > 1$). Quello che si ottiene è sostanzialmente il modello generale visto in sezione 3.5.1 in cui però si pone $\Delta E_0 = \{0\}$. La situazione è illustrata in figura 3.7.

Ancora una volta lo stato 0 indica assenza di radiazione solare e quindi di energia per ricaricare la batteria. Gli stati $\{1, 2, \dots, X\}$ rappresentano, invece, livelli via via crescenti di radiazione.

3.6 Radio, sensore, unità di controllo e buffer

Si sono già evidenziate le due attività fondamentali che si suppone siano eseguite dal sensore radio: la ricezione e ritrasmissione di pacchetti esogeni e la rilevazione di un particolare fenomeno d'interesse con invio dei dati raccolti agli altri nodi della rete. Tali operazioni prevedono l'utilizzo della radio, del sensore, dell'unità di controllo e

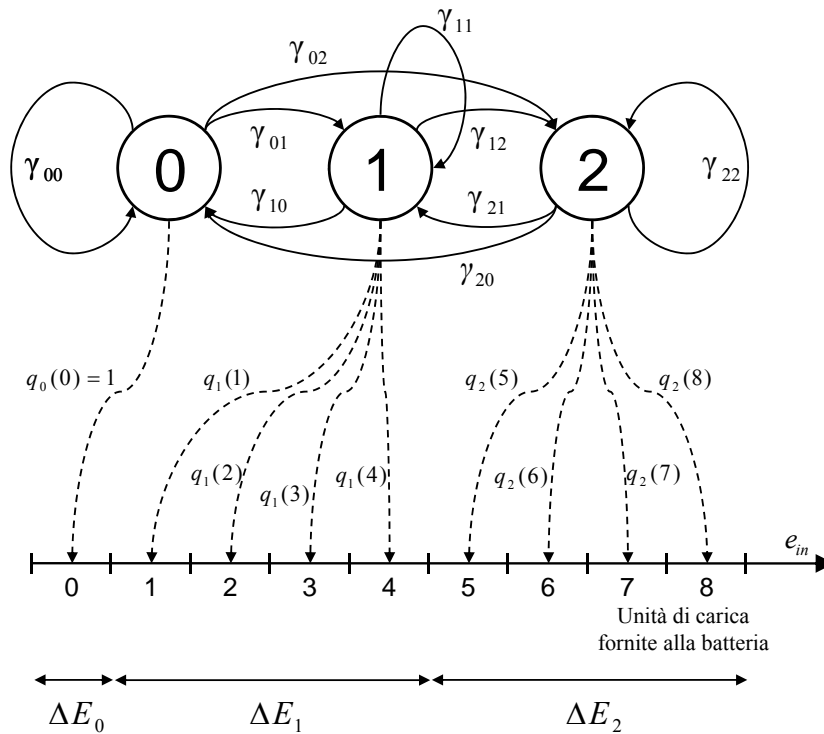


Figura 3.7 Modello per l'energia solare a più stati con $X = 2$, $E_{MAX} = 8$, $\Delta E_0 = \{0\}$, $\Delta E_1 = \{1, 2, 3, 4\}$ e $\Delta E_2 = \{5, 6, 7, 8\}$

dei buffer per la memorizzazione temporanea delle informazioni. Si propone a questo punto un modello in grado di rappresentare il funzionamento di tali componenti.

Si consideri allora un generico slot temporale. Si assume che la radio possa essere in uno tra tre stati ammissibili: *ricezione*, *trasmissione*, *idle* (radio spenta). Se nello slot corrente la radio è in trasmissione avviene l'invio di uno dei pacchetti presenti nei buffer, scelto in base alla politica specificata dall'unità di controllo; se, invece, la radio è in ricezione viene ascoltato il canale al fine di rilevare l'eventuale presenza di traffico esogeno e, in caso affermativo, viene ricevuto un pacchetto dati e memorizzato nel buffer dedicato; infine, se la radio è in idle non viene eseguita né la ricezione né la trasmissione di alcun pacchetto.

Il sensore, invece, in ciascuno slot può essere *attivo* o *spento*. Se è attivo avviene la rilevazione del fenomeno ambientale che porta alla memorizzazione di un nuovo pacchetto dati nel buffer solo se le informazioni acquisite sono aggiornate rispetto alle precedenti; se il sensore è spento, invece, non avviene alcuna rilevazione.

Si noti che nel caso in cui in uno slot sia attiva la radio oppure il sensore, necessariamente dovrà essere attiva anche l'unità di controllo per coordinare le operazioni. Inoltre, ricordando le assunzioni di sezione 3.2, la durata di ogni slot assicura che ognuna delle azioni appena citate possa essere eseguita completamente.

In uno slot, sia quando la radio è in ricezione che quando l'unità di sensing compie una misurazione, non è detto che l'operazione termini con successo, cioè con la memorizzazione di un pacchetto dati nel buffer dedicato. Si definiscono allora:

- P_{rx} = probabilità di arrivo di un pacchetto esogeno per slot.
- P_{sx} = probabilità che l'attività di sensing in uno slot porti all'acquisizione di informazioni aggiornate.

Allora, se in uno slot la radio è in ricezione, con probabilità P_{rx} verrà ricevuto e memorizzato nel buffer opportuno un nuovo pacchetto. Allo stesso modo se il sensore è attivo, con probabilità P_{sx} si otterrà un nuovo pacchetto da memorizzare. I due parametri P_{rx} e P_{sx} sono molto importanti: variare il valore di P_{rx} consente di rappresentare situazioni in cui il traffico esogeno è di diversa intensità; analogamente P_{sx} permette di modellare la frequenza con cui varia il fenomeno che interessa monitorare. Valori elevati di P_{rx} e P_{sx} , ad esempio, rappresentano un traffico in entrata al sensore radio molto intenso e un fenomeno fisico che muta molto velocemente.

Le operazioni svolte utilizzando la radio e l'unità di sensing richiedono naturalmente una certa quantità di energia che viene fornita dalla batteria. In particolare si definiscono:

- E_{id} = numero di unità di carica per slot richieste dalla radio in idle.
- E_{tx} = numero di unità di carica per slot richieste per la trasmissione.
- E_{rx} = numero di unità di carica per slot richieste per la ricezione.
- E_{sx} = numero di unità di carica per slot richieste dall'attività di sensing.

Ricordando ancora le assunzioni di sezione 3.2, le quantità appena definite rappresentano l'effettivo dispendio energetico delle varie operazioni, ovvero comprendono i consumi di tutti gli elementi circuitali coinvolti nelle diverse attività (come ad esempio l'unità di controllo).

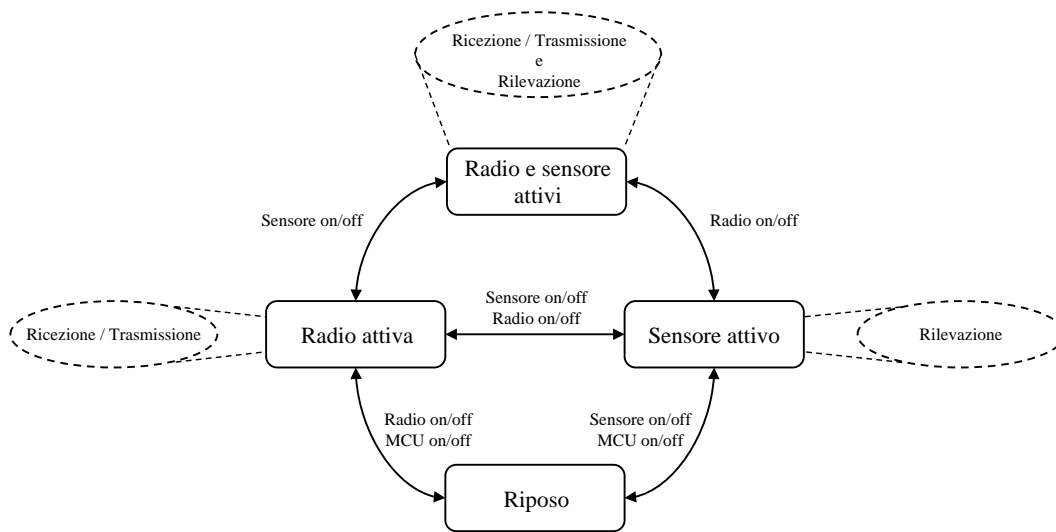


Figura 3.8 Modello per radio e sensore indipendenti.

Rimane da stabilire a questo punto se è ammissibile o meno che in uno stesso slot la radio e l'unità di sensing operino contemporaneamente. Questo dipende dal tipo di sensore radio considerato e da come sono gestite le attività di sensing e relay. Vale la pena, allora, considerare separatamente entrambi i casi.

3.6.1 Radio e sensore indipendenti

Nel caso in cui la radio e il sensore operino in modo indipendente è possibile che nello stesso slot entrambe le componenti siano attive. Dunque è ammissibile che la radio esegua una determinata operazione, come la ricezione o la trasmissione di un pacchetto, e che contemporaneamente nello stesso slot il sensore sia attivo per compiere la rilevazione. Quello che si ottiene è mostrato in figura 3.8 che riprende l'analogo modello proposto in [30].

In ogni slot, dunque, si presentano quattro possibilità:

- *Riposo*: sia la radio che il sensore sono spenti (così come l'unità di controllo) e non viene svolta nessuna attività.
- *Radio attiva*: la radio è accesa in ricezione o trasmissione mentre il sensore è inattivo.

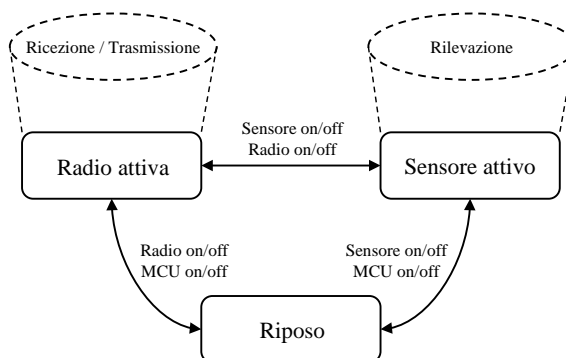


Figura 3.9 Modello per radio e sensore mutuamente esclusivi.

- *Sensore attivo*: la radio è in idle mentre il sensore compie la rilevazione.
- *Radio e sensore attivi*: entrambe le componenti sono attive e svolgono la loro attività in modo indipendente.

3.6.2 Radio e sensore mutuamente esclusivi

In alcuni casi è possibile che la radio e il sensore non possano essere utilizzati contemporaneamente per svolgere le rispettive attività; in un determinato slot, quindi, può essere attiva al più una delle due componenti. La situazione è schematizzata in figura 3.9: non è ammesso che nello stesso slot la radio sia attiva mentre viene eseguita la rilevazione e, viceversa, non è possibile il sensing se la radio è accesa per ricevere o trasmettere un pacchetto.

3.6.3 Buffer

I buffer per il traffico locale e per quello esogeno sono coinvolti nelle operazioni svolte dal sensore e dalla radio in quanto consentono la memorizzazione dei pacchetti in attesa che questi vengano trasmessi. In particolare, un pacchetto ricevuto dalla radio viene salvato nel buffer per il traffico esogeno, mentre un pacchetto originato dal sensore viene memorizzato nel buffer per il traffico locale. Si definiscono allora:

- W = dimensione del buffer per il traffico locale (numero massimo di pacchetti memorizzabili).

- R = dimensione del buffer per il traffico esogeno.

All'inizio di un generico slot la situazione dei buffer è rappresentata dalla coppia di variabili (w, r) dove $w \in \{0, 1, \dots, W\}$ indica il numero di pacchetti nel buffer per il traffico locale mentre $r \in \{0, 1, \dots, R\}$ specifica il numero di pacchetti presenti nel buffer per il traffico esogeno.

3.6.4 Modalità di funzionamento

In base agli obiettivi discussi in sezione 3.3, vi è l'esigenza di definire un parametro che consenta di specificare, per il sensore radio considerato, quale attività tra inoltro e monitoraggio abbia maggior rilevanza e in che misura. Questo è necessario poiché ciascun nodo di una WSN ricopre un proprio ruolo all'interno della rete il quale può richiedere, ad esempio, una maggior attività di sensing piuttosto che di inoltro o viceversa.

Con questo intento si introduce una variabile di stato binaria, indicata con v , che in ciascuno slot può assumere solo i valori 0 o 1. Tale variabile indica la *modalità di funzionamento* del sensore radio nello slot corrente: se $v = 0$ il dispositivo opera dando precedenza all'attività di rilevazione del fenomeno ambientale, se $v = 1$, invece, viene prediletta l'attività di relay dei pacchetti esogeni.

In ogni slot il valore assunto da v è determinato dalla distribuzione $d_v(\cdot)$ così definita:

$$d_v(\bar{v}) = P[v = \bar{v}] = \begin{cases} \beta & \text{se } \bar{v} = 0 \\ 1 - \beta & \text{se } \bar{v} = 1 \end{cases}, \quad \beta \in [0, 1]. \quad (3.5)$$

Secondo la (3.5), β è la probabilità che in un qualsiasi slot v sia pari a 0; dunque β può essere interpretata come la frazione di tempo durante la quale il sensore radio predilige l'attività di monitoraggio del fenomeno d'interesse.

Alla luce di quanto detto β è proprio il parametro che si mirava ad ottenere. Variando il valore di β , infatti, è possibile modulare a piacere l'importanza per il sensore radio di ciascuna delle due attività: se $\beta > 0.5$ ha maggiore rilevanza l'attività di sensing, se $\beta < 0.5$ quella di inoltro; quanto più β è vicina ai valori limite 1 o 0 tanto più spiccata è la preferenza per l'una o l'altra attività.

3.7 Modello complessivo per il sensore radio

I modelli presentati nelle sezioni precedenti per le singole componenti del sensore radio possono essere combinati tra loro per ottenere il modello complessivo dell'intero dispositivo. Ciò che si propone per rappresentare il sensore radio è un processo decisionale di Markov con istanti di decisione, stati, azioni, probabilità di transizione e costi che vengono ora definiti in modo opportuno.

3.7.1 Istanti di decisione

In base alle assunzioni di sezione 3.2, si suppone che il tempo sia suddiviso e misurato in slot di durata tale da consentire la completa ricezione o trasmissione o rilevazione di un pacchetto dati; il sensore radio è quindi rappresentato come un sistema a tempo discreto.

Si assume, allora, che gli istanti di decisione, compresi nell'insieme \mathcal{T} , coincidano con l'inizio di ciascuno slot temporale. Poiché il sensore radio è in grado di operare per un tempo potenzialmente illimitato, grazie alla fonte di energia rinnovabile che ricarica la batteria, l'insieme \mathcal{T} è infinito e dunque del tipo $\mathcal{T} = \{0, 1, 2, \dots\}$. Il processo decisionale è quindi ad orizzonte infinito.

3.7.2 Stati

In corrispondenza di ciascun istante di decisione, cioè all'inizio di ogni slot, il sensore radio occupa un determinato stato. L'insieme di tutti i possibili stati del sistema rimane invariato nel tempo e dunque è indicato semplicemente con \mathcal{S} .

Un qualsiasi stato $i \in \mathcal{S}$ ha la forma

$$i = (n, x, w, r, v), \quad (3.6)$$

dove le varie componenti sono le variabili di stato, già definite, che descrivono la condizione in cui si trovano gli elementi costitutivi del sensore radio. Per chiarezza se ne ricorda di seguito il significato:

- n = numero di unità di carica rimaste nella batteria e quindi disponibili nello slot corrente per le attività del dispositivo, $n \in \{0, 1, \dots, N\}$.

- x = stato occupato dalla fonte di energia rinnovabile nello slot appena concluso. Indica l'intensità del flusso di energia che ha ricaricato la batteria; $x \in \{0, 1, \dots, X\}$.
- w = numero di pacchetti memorizzati nel buffer per il traffico locale in attesa di essere trasmessi, $w \in \{0, 1, \dots, W\}$.
- r = numero di pacchetti memorizzati nel buffer per il traffico esogeno in attesa di essere trasmessi, $r \in \{0, 1, \dots, R\}$.
- v = modalità di funzionamento del sensore radio nello slot corrente, $v \in \{0, 1\}$.

Il numero totale di stati del processo decisionale è dunque

$$S = |\mathcal{S}| = (N + 1)(X + 1)(W + 1)(R + 1) \cdot 2. \quad (3.7)$$

Assumendo che inizialmente la batteria sia completamente carica, lo stato iniziale che rappresenta la condizione di partenza del sensore radio può essere $s_0 = (N, 0, 0, 0, v_0)$ oppure $s_0 = (N, x_0, 0, 0, v_0)$ con x_0 determinato in base alla distribuzione stazionaria della catena di Markov rappresentativa della fonte di energia rinnovabile e v_0 stabilito secondo la distribuzione $d_v(\cdot)$ definita in (3.5).

3.7.3 Azioni

Considerato un generico istante di decisione $t \in \mathcal{T}$ e indicato con $i \in \mathcal{S}$ lo stato occupato dal sistema, l'unità di controllo ha la responsabilità di stabilire quale azione debba essere compiuta dal sensore radio nello slot corrente. Le azioni che possono essere eseguite dipendono dallo stato i e in particolare dal numero di unità di carica rimaste nella batteria e dalla situazione dei buffer. Inoltre è necessario distinguere il caso con radio e sensore indipendenti da quello con radio e sensore mutuamente esclusivi.

A Radio e sensore indipendenti

In questo caso in qualsiasi slot temporale il sensore radio può eseguire o meno l'operazione di sensing e, contestualmente e indipendentemente dall'attività di rilevazione,

può mantenere la radio spenta oppure attivarla per la trasmissione o ricezione di un pacchetto.

In ogni slot, quindi, il dispositivo compie un'azione del tipo

$$a = (a_1, a_2).$$

La componente $a_1 \in \{ON, OFF\}$ specifica se nello slot considerato il sensore è attivo (ON) o spento (OFF). La componente $a_2 \in \{ID, RX, TX_w, TX_r\}$ indica, invece, lo stato della radio: può essere in idle (ID), ricezione (RX), trasmissione di un pacchetto del buffer per il traffico locale (TX_w) oppure del buffer per il traffico esogeno (TX_r).

Dunque l'insieme \mathcal{A} , indipendente da t , di tutte le possibili azioni che il sensore radio può eseguire in un dato slot è

$$\begin{aligned} \mathcal{A} = \{ & (OFF, ID), (OFF, RX), (OFF, TX_w), (OFF, TX_r), \\ & (ON, ID), (ON, RX), (ON, TX_w), (ON, TX_r) \}. \end{aligned} \quad (3.8)$$

Per ogni stato $i \in \mathcal{S}$, poi, l'action set $\mathcal{A}(i) \subseteq \mathcal{A}$ è determinato in base alla situazione della batteria, cioè al valore di n , e dei buffer, quindi ai valori di w e r , secondo le seguenti regole:

- *Regola 1:* un'azione a non può appartenere ad $\mathcal{A}(i)$ se lo stato della batteria in i non consente di fornire al sensore radio un numero sufficiente di unità di carica per compiere l'azione a , ovvero se le unità disponibili n sono meno del necessario.
- *Regola 2:* se nello stato i il buffer per il traffico locale è vuoto ($w = 0$), un'azione a che preveda la trasmissione di un pacchetto originato dal sensore non può appartenere ad $\mathcal{A}(i)$ dato che non c'è nulla da trasmettere. Allo stesso modo se il buffer per il traffico esterno è vuoto ($r = 0$) non è possibile trasmettere alcun pacchetto esogeno.
- *Regola 3:* se nello stato i il buffer per il traffico locale è pieno ($w = W$), un'azione a che preveda di eseguire una nuova rilevazione non può appartenere ad $\mathcal{A}(i)$ dato che non c'è posto per memorizzare un eventuale nuovo pacchetto. Allo stesso modo se il buffer per il traffico esogeno è pieno ($r = R$) non è possibile ricevere nuovi pacchetti dall'esterno.

In base alle osservazioni precedenti si può utilizzare una funzione ausiliaria $f_A(i, a)$ per definire in modo compatto l'action set $\mathcal{A}(i)$ di ogni $i \in \mathcal{S}$. Nello specifico si pone

$$\mathcal{A}(i) = \{a \in \mathcal{A} : f_A(i, a) = 1\},$$

dove, per ogni stato $i = (n, x, w, r, v) \in \mathcal{S}$ e per ogni azione $a = (a_1, a_2) \in \mathcal{A}$, la funzione $f_A(i, a)$ è definita come

$$f_A(i, a) = \begin{cases} 0 & \begin{cases} n < E_{id} \text{ e } a_2 = ID \\ n < E_{rx} \text{ e } a_2 = RX \\ n < E_{tx} \text{ e } (a_2 = TX_w \text{ o } a_2 = TX_r) \\ n < E_{id} + E_{sx} \text{ e } a_1 = ON \text{ e } a_2 = ID \\ n < E_{rx} + E_{sx} \text{ e } a_1 = ON \text{ e } a_2 = RX \\ n < E_{tx} + E_{sx} \text{ e } a_1 = ON \text{ e } (a_2 = TX_w \text{ o } a_2 = TX_r) \\ w = 0 \text{ e } a_2 = TX_w \\ w = W \text{ e } a_1 = ON \\ r = 0 \text{ e } a_2 = TX_r \\ r = R \text{ e } a_2 = RX \end{cases} \\ 1 & \text{altrimenti} \end{cases} \quad (3.9)$$

B Radio e sensore mutuamente esclusivi

Se la radio e il sensore sono mutuamente esclusivi le due componenti non possono essere contemporaneamente attive in uno stesso slot. Conseguentemente l'insieme \mathcal{A} di tutte le possibili azioni si riduce a

$$\mathcal{A} = \{ID, RX, TX_w, TX_r, SX\}. \quad (3.10)$$

Quindi in ciascuno slot il sensore radio compie un'azione $a \in \mathcal{A}$ scelta tra: mantenere sia la radio che il sensore spenti (ID), ricevere un pacchetto esogeno (RX), trasmettere un pacchetto memorizzato nel buffer per il traffico locale (TX_w), trasmettere un pacchetto prelevato dal buffer per il traffico esogeno (TX_r) ed eseguire una nuova rilevazione (SX).

Le azioni contenute nell'action set $\mathcal{A}(i) \subseteq \mathcal{A}$ di ogni stato $i \in \mathcal{S}$ sono determinate mediante le stesse tre regole viste nel caso con radio e sensore indipendenti.

Ancora una volta, per ogni $i \in \mathcal{S}$, l'insieme $\mathcal{A}(i)$ può essere definito in modo compatto ricorrendo ad una funzione ausiliaria $f_B(i, a)$. Si pone cioè

$$\mathcal{A}(i) = \{a \in \mathcal{A} : f_B(i, a) = 1\},$$

dove per ogni stato $i = (n, x, w, r, v) \in \mathcal{S}$ e per ogni azione $a \in \mathcal{A}$ la funzione $f_B(i, a)$ è definita come

$$f_B(i, a) = \begin{cases} 0 & \begin{aligned} & n < E_{id} \text{ e } a = ID \\ & n < E_{rx} \text{ e } a = RX \\ & n < E_{tx} \text{ e } (a = TX_w \text{ o } a = TX_r) \\ & n < E_{id} + E_{sx} \text{ e } a = SX \\ & w = 0 \text{ e } a = TX_w \\ & w = W \text{ e } a = SX \\ & r = 0 \text{ e } a = TX_r \\ & r = R \text{ e } a = RX \end{aligned} \\ 1 & \text{altrimenti} \end{cases} \quad (3.11)$$

3.7.4 Probabilità di transizione

Si supponga di indicare con $i = (n, x, w, r, v) \in \mathcal{S}$ lo stato del sensore radio in un determinato istante di decisione, corrispondente all'inizio di un particolare slot temporale.

I valori w e r (insieme ad n) determinano l'action set $\mathcal{A}(i)$, come visto in sezione 3.7.3, e vengono aggiornati in base all'azione $a \in \mathcal{A}(i)$ che viene effettivamente eseguita nello slot corrente. In particolare il valore di w (il valore di r) viene decrementato di un'unità se avviene la trasmissione di un pacchetto memorizzato nel buffer per il traffico locale (esogeno), viene incrementato di un'unità se viene eseguita con successo una nuova rilevazione (ricezione) mentre rimane inalterato negli altri casi.

Con x è indicato, invece, lo stato occupato dalla fonte di energia rinnovabile nello slot appena concluso. L'apporto energetico che sarà dato alla batteria nello slot corrente è quindi determinato dall'evoluzione della catena di Markov rappresentativa della fonte

rinnovabile descritta in sezione 3.5: ci sarà una transizione da x a x' e conseguente acquisizione di e_{in} unità di carica da parte della batteria con probabilità $\gamma_{xx'} q_{x'}(e_{in})$. Inoltre, x' sarà il nuovo valore che verrà considerato all'inizio dello slot successivo.

La componente v , poi, specifica la modalità di funzionamento del sensore radio nello slot corrente. Il valore di v si aggiorna ad ogni slot in accordo con la distribuzione $d_v(\cdot)$ definita in (3.5).

Infine, il valore di n , che indica le unità di carica residue nella batteria e disponibili per lo slot corrente, viene aggiornato tenendo conto dell'apporto energetico e_{in} fornito dalla fonte rinnovabile e del consumo di energia e_{out} dovuto alla particolare azione eseguita nello slot.

Definite le due funzioni di supporto

$$\delta_a^{\vartheta} = \begin{cases} 1 & \text{se } a = \vartheta \\ 0 & \text{altrimenti,} \end{cases}$$

$$\delta_{a_1, a_2}^{\vartheta_1, \vartheta_2} = \begin{cases} 1 & \text{se } a_1 = \vartheta_1 \text{ e } a_2 = \vartheta_2 \\ 0 & \text{altrimenti,} \end{cases}$$

è possibile fornire una formulazione compatta per il numero di unità di carica $e_{out}(a)$ richieste nello slot corrente dal sensore radio in funzione dell'azione a eseguita. In particolare, se si considera il caso A di sezione 3.7.3 con radio e sensore indipendenti vale la relazione

$$e_{out}(a) = e_{out}(a_1, a_2) = \delta_{a_2}^{ID} E_{id} + \delta_{a_2}^{RX} E_{rx} + \delta_{a_2}^{TX_w} E_{tx} + \delta_{a_2}^{TX_r} E_{tx} + \delta_{a_1}^{ON} E_{sx},$$

mentre nel caso B con radio e sensore mutuamente esclusivi vale

$$e_{out}(a) = \delta_a^{ID} E_{id} + \delta_a^{RX} E_{rx} + \delta_a^{TX_w} E_{tx} + \delta_a^{TX_r} E_{tx} + \delta_a^{SX} E_{sx}.$$

In conclusione, indicato con $i \in \mathcal{S}$ lo stato del sensore radio all'inizio di un generico slot, il dispositivo può eseguire in tale slot una delle azioni $a \in \mathcal{A}(i)$; in base all'azione scelta diversi sono gli stati verso cui il sistema può transitare.

Si consideri in primo luogo il caso che prevede la radio e il sensore indipendenti: dato lo stato corrente $i = (n, x, w, r, v)$, in tabella 3.1 sono riportate le possibili transizioni,

$p_{ij}(a_1, a_2)$	$j = (n', x', w', r', v')$
$\{\delta_{a_1, a_2}^{ON, RX}[P_{rx}P_{sx}]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a_1, a_2), x', w + 1, r + 1, \bar{v})$
$\{\delta_{a_1, a_2}^{ON, ID}[P_{sx}] + \delta_{a_1, a_2}^{ON, RX}[(1 - P_{rx})P_{sx}]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a_1, a_2), x', w + 1, r, \bar{v})$
$\{\delta_{a_1, a_2}^{OFF, RX}[P_{rx}] + \delta_{a_1, a_2}^{ON, RX}[P_{rx}(1 - P_{sx})]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a_1, a_2), x', w, r + 1, \bar{v})$
$\{\delta_{a_1, a_2}^{OFF, ID} + \delta_{a_1, a_2}^{ON, RX}[(1 - P_{rx})(1 - P_{sx})] + \delta_{a_1, a_2}^{ON, TX_w}[P_{sx}] + \delta_{a_1, a_2}^{ON, ID}[1 - P_{sx}] + \delta_{a_1, a_2}^{OFF, RX}[1 - P_{rx}]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a_1, a_2), x', w, r, \bar{v})$
$\{\delta_{a_1, a_2}^{OFF, TX_w} + \delta_{a_1, a_2}^{ON, TX_w}[(1 - P_{sx})]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a_1, a_2), x', w - 1, r, \bar{v})$
$\{\delta_{a_1, a_2}^{OFF, TX_r} + \delta_{a_1, a_2}^{ON, TX_r}[(1 - P_{sx})]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a_1, a_2), x', w, r - 1, \bar{v})$
$\{\delta_{a_1, a_2}^{ON, TX_r}[P_{sx}]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a_1, a_2), x', w + 1, r - 1, \bar{v})$

Tabella 3.1 Probabilità di transizione con radio e sensore indipendenti.

le probabilità di transizione e il relativo stato d'arrivo $j = (n', x', w', r', v')$ in funzione dell'azione $a = (a_1, a_2)$ scelta.

Nella tabella 3.2, invece, si fa riferimento ad un dispositivo con radio e sensore mutuamente esclusivi: dato lo stato corrente $i = (n, x, w, r, v)$, sono indicate le transizioni possibili, le relative probabilità di transizione e lo stato d'arrivo $j = (n', x', w', r', v')$ in funzione dell'azione a scelta.

3.7.5 Costi

I costi accumulati nel singolo slot devono essere definiti opportunamente alla luce degli obiettivi discussi in sezione 3.3. In particolare, l'obiettivo generale può essere

$p_{ij}(\mathbf{a})$	$\mathbf{j} = (n', x', w', r', v')$
$\{\delta_a^{SX} [P_{sx}]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a), x', w + 1, r, \bar{v})$
$\{\delta_a^{RX} [P_{rx}]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a), x', w, r + 1, \bar{v})$
$\{\delta_a^{ID} + \delta_a^{RX} [1 - P_{rx}] + \delta_a^{SX} [1 - P_{sx}]\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a), x', w, r, \bar{v})$
$\{\delta_a^{TXw}\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a), x', w - 1, r, \bar{v})$
$\{\delta_a^{TXr}\} \gamma_{xx'} q_{x'}(e_{in}) d_v(\bar{v})$	$(n + e_{in} - e_{out}(a), x', w, r - 1, \bar{v})$

Tabella 3.2 Probabilità di transizione con radio e sensore mutuamente esclusivi.

sintetizzato come di seguito.

Obiettivo *Determinare la politica di scheduling ottima che, una volta applicata dall'unità di controllo, porti il sensore radio a massimizzare il throughput, ovvero il numero medio di pacchetti trasmessi per slot, rispettando il vincolo sul livello di guardia della batteria.*

Si vuole, inoltre, che il dispositivo possa dare più rilevanza ad una tra le due attività di monitoraggio ed inoltro e che tale rilevanza sia più o meno marcata in base alla scelta di un parametro. Si è già visto che a tal fine il parametro adatto su cui agire è β il quale determina la distribuzione di v in ogni slot.

L'idea è di individuare la politica ottima adottando il criterio di ottimizzazione del costo medio per slot. Vi è quindi la necessità di definire i costi sul singolo slot in modo tale che, una volta determinata la politica ottima secondo il criterio del costo medio per slot, essa sia di conseguenza anche la politica che porta il sensore radio a massimizzare il throughput. Allora appare evidente che in ogni slot dovrà essere associato un costo basso alle azioni che prevedono la trasmissione di un pacchetto, mentre dovrà essere

assegnato un costo alto alle azioni che non prevedono nessuna attività da parte del sensore radio. In questo modo la politica che minimizza il costo medio per slot sarà quella che porta al maggior numero di trasmissioni per slot e dunque sarà proprio la politica che spinge il dispositivo a massimizzare il numero di pacchetti trasmessi.

Inoltre, in base a quanto detto in sezione 3.6.4, negli slot in cui $v = 0$ ha precedenza l'attività di monitoraggio del fenomeno d'interesse e si vuole, quindi, che il sensore radio operi in modo da massimizzare il throughput dei pacchetti generati dall'attività di sensing. Viceversa negli slot in cui $v = 1$ l'interesse è che il dispositivo massimizzi il throughput dei pacchetti esogeni.

Per definire in modo preciso i costi sul singolo slot si introducono:

- C_{tx} = costo per slot in caso di mancata trasmissione di un pacchetto,
- C_{px} = costo per slot in caso di mancata acquisizione di un nuovo pacchetto,
- C_{plus} = costo per slot supplementare se il livello della batteria è sotto il livello di guardia,

dove si assume che $C_{tx} > C_{px}$ e che tutti i costi siano superiormente limitati.

Si consideri, ora, uno slot in cui $v = 0$. In base alle osservazioni precedenti, si attribuisce il costo C_{tx} alle azioni che non prevedono la trasmissione di un pacchetto memorizzato nel buffer per il traffico locale; in più si assegna un costo C_{px} alle azioni per le quali non viene eseguito il sensing. Si procede, poi, in modo speculare nel caso $v = 1$: il costo C_{tx} è associato alle azioni che non prevedono la trasmissione di un pacchetto memorizzato nel buffer per il traffico esogeno e il costo C_{px} viene attribuito alle azioni per le quali non avviene la ricezione di un nuovo pacchetto. Il costo C_{px} ha il ruolo di rendere più appetibili le azioni nelle quali si cerca di acquisire un nuovo pacchetto penalizzando le altre; memorizzare nuovi pacchetti, infatti, consente la trasmissione negli slot successivi. Infine, negli slot in cui n è al di sotto del livello di guardia N_{lim} , a tutte le azioni è attribuito un costo supplementare C_{plus} . Così facendo si introduce una penalizzazione quando il livello di carica della batteria è troppo basso.

Se i costi vengono definiti secondo il procedimento appena descritto, negli slot in cui $v = 0$ l'unità di controllo accumulerà il costo minimo se l'azione eseguita dal sensore

$a = (a_1, a_2)$	$g(i, a)$			
	$v = 0 \text{ e } n \geq N_{lim}$	$v = 1 \text{ e } n \geq N_{lim}$	$v = 0 \text{ e } n < N_{lim}$	$v = 1 \text{ e } n < N_{lim}$
(OFF, ID)	$C_{tx} + C_{px}$	$C_{tx} + C_{px}$	$C_{tx} + C_{px} + C_{plus}$	$C_{tx} + C_{px} + C_{plus}$
(OFF, RX)	$C_{tx} + C_{px}$	C_{tx}	$C_{tx} + C_{px} + C_{plus}$	$C_{tx} + C_{plus}$
(OFF, TX _w)	C_{px}	$C_{tx} + C_{px}$	$C_{px} + C_{plus}$	$C_{tx} + C_{px} + C_{plus}$
(OFF, TX _r)	$C_{tx} + C_{px}$	C_{px}	$C_{tx} + C_{px} + C_{plus}$	$C_{px} + C_{plus}$
(ON, ID)	C_{tx}	$C_{tx} + C_{px}$	$C_{tx} + C_{plus}$	$C_{tx} + C_{px} + C_{plus}$
(ON, RX)	C_{tx}	C_{tx}	$C_{tx} + C_{plus}$	$C_{tx} + C_{plus}$
(ON, TX _w)	0	$C_{tx} + C_{px}$	C_{plus}	$C_{tx} + C_{px} + C_{plus}$
(ON, TX _r)	C_{tx}	C_{px}	$C_{tx} + C_{plus}$	$C_{px} + C_{plus}$

Tabella 3.3 Costi per slot con radio e sensore indipendenti.

radio prevede la trasmissione di un pacchetto di sensing, mentre negli slot in cui $v = 1$ le azioni più convenienti saranno quelle che implicano la trasmissione di un pacchetto esogeno. Ecco che allora la politica che minimizza il costo medio per slot in questo modo è anche quella che porta a massimizzare il throughput.

Si può procedere, a questo punto, formalizzando i concetti introdotti finora. Si consideri innanzitutto il caso con radio e sensore indipendenti: sia $i = (n, x, w, r, v)$ lo stato occupato dal sistema all'inizio di un generico slot e sia $a = (a_1, a_2)$ l'azione compiuta dal sensore radio. Il costo $g(i, a)$ accumulato nello slot è allora

$$g(i, a) = C_{tx} + C_{px} - \delta_v^0 \delta_{a_2}^{TX_w} C_{tx} - \delta_v^1 \delta_{a_2}^{TX_r} C_{tx} - \delta_v^0 \delta_{a_1}^{ON} C_{px} - \delta_v^1 \delta_{a_2}^{RX} C_{px} + \eta_n^{N_{lim}} C_{plus},$$

dove

$$\eta_n^L = \begin{cases} 1 & \text{se } n < L \\ 0 & \text{altrimenti} \end{cases}$$

In tabella 3.3 sono riportati i valori di $g(i, a)$ nei vari casi e per le diverse azioni.

Se, invece, la radio e il sensore sono mutuamente esclusivi, il costo $g(i, a)$ accumulato in uno slot in cui lo stato è $i = (n, x, w, r, v)$ e l'azione compiuta è a risulta

$$g(i, a) = C_{tx} + C_{px} - \delta_v^0 \delta_a^{TX_w} C_{tx} - \delta_v^1 \delta_a^{TX_r} C_{tx} - \delta_v^0 \delta_a^{SX} C_{px} - \delta_v^1 \delta_a^{RX} C_{px} + \eta_n^{N_{lim}} C_{plus}.$$

I valori di $g(i, a)$ per le diverse azioni e nei vari casi sono mostrati nella tabella 3.4.

a	$g(i, a)$			
	$v = 0 \text{ e } n \geq N_{lim}$	$v = 1 \text{ e } n \geq N_{lim}$	$v = 0 \text{ e } n < N_{lim}$	$v = 1 \text{ e } n < N_{lim}$
ID	$C_{tx} + C_{px}$	$C_{tx} + C_{px}$	$C_{tx} + C_{px} + C_{plus}$	$C_{tx} + C_{px} + C_{plus}$
RX	$C_{tx} + C_{px}$	C_{tx}	$C_{tx} + C_{px} + C_{plus}$	$C_{tx} + C_{plus}$
TX_w	C_{px}	$C_{tx} + C_{px}$	$C_{px} + C_{plus}$	$C_{tx} + C_{px} + C_{plus}$
TX_r	$C_{tx} + C_{px}$	C_{px}	$C_{tx} + C_{px} + C_{plus}$	$C_{px} + C_{plus}$
SX	C_{tx}	$C_{tx} + C_{px}$	$C_{tx} + C_{plus}$	$C_{tx} + C_{px} + C_{plus}$

Tabella 3.4 Costi per slot con radio e sensore mutuamente esclusivi.

3.8 Osservazioni conclusive sul modello

In sezione 3.7 sono stati definiti tutti gli elementi del processo decisionale di Markov proposto per rappresentare le componenti principali e il funzionamento di un sensore radio alimentato mediante una fonte di energia rinnovabile.

L'aspetto importante da sottolineare è che tale processo decisionale di Markov soddisfa le assunzioni poste in sezione 2.3.1: è a tempo discreto, dato che il tempo è suddiviso in slot temporali; ad orizzonte infinito, visto che l'evoluzione del dispositivo ha una durata potenzialmente illimitata; stazionario, in quanto lo spazio degli stati, gli action set, le probabilità di transizione e i costi sono invarianti nel tempo; l'insieme degli stati e gli action set sono finiti e i costi superiormente limitati.

Il fatto che siano soddisfatte le assunzioni di sezione 2.3.1 garantisce che, una volta applicato al processo il criterio di ottimizzazione del costo medio per slot, siano validi i risultati visti in sezione 2.4. In particolare è possibile fare riferimento alle equazioni di ottimalità (2.47) e (2.48) per determinare il costo medio per slot ottimo e la relativa politica (stazionaria) ottima.

Per come sono stati definiti i costi sul singolo slot in sezione 3.7.5, determinare la politica ottima secondo il criterio del costo medio per slot significa individuare quale sia la politica che l'unità di controllo deve adottare per portare il sensore radio a massimizzare il throughput. Allora, una volta stabilito come risolvere le equazioni (2.47) e (2.48), si è in grado di calcolare al variare dei parametri in gioco quale sia la politica (di scheduling) che porti il dispositivo ad operare in modo ottimo secondo il criterio del costo medio per slot e quindi a massimizzare le prestazioni.

4

Algoritmo di ottimizzazione

Dato un problema decisionale di Markov con il criterio di ottimizzazione del costo medio per slot, le possibili strategie utilizzabili per determinare il costo medio ottimo e la relativa politica ottima sono diverse a seconda delle caratteristiche strutturali del processo decisionale considerato. Per questo motivo all'inizio del capitolo si introduce un'utile classificazione dei processi decisionali di Markov che consenta di individuare quali siano le proprietà del modello proposto per il sensore radio nel capitolo 3. In base alle caratteristiche del modello si propone un opportuno algoritmo di ottimizzazione che, una volta implementato, permetta di calcolare le politiche ottime per il sensore radio.

4.1 Classificazione dei processi decisionali di Markov

Nel capitolo 2, in particolare con la proposizione 2.4.8, si è giunti a dimostrare che per un qualsiasi processo decisionale di Markov, per cui si assume siano soddisfatte le assunzioni di sezione 2.3.1, esiste sempre una politica stazionaria ottima secondo il criterio del costo medio per slot. Inoltre con la definizione 2.4.2 si sono introdotte le equazioni di ottimalità che consentono di individuare il costo medio ottimo e la corrispondente politica stazionaria ottima; tali equazione sono

$$J(i) = \min_{a \in \mathcal{A}(i)} \sum_{j \in \mathcal{S}} p_{ij}(a) J(j), \quad i \in \mathcal{S}, \quad (4.1)$$

e

$$J(i) + h(i) = \min_{a \in \hat{\mathcal{A}}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h(j) \right], \quad i \in \mathcal{S}, \quad (4.2)$$

dove, per ogni i , $\hat{\mathcal{A}}(i)$ rappresenta l'insieme delle azioni che realizzano il minimo in (4.1).

Le tecniche che possono essere adottate per risolvere le equazioni (4.1) e (4.2) cambiano in base alle caratteristiche del processo decisionale di Markov considerato.

Nello specifico è molto rilevante la struttura delle catene di Markov indotte nel processo dalle politiche stazionarie.

Si definiscono allora alcune classi a cui appartengono processi decisionali di Markov con particolari e diverse proprietà, si individuano le relazioni che intercorrono tra tali classi e si evidenzia in che modo possa essere classificato un dato processo decisionale [2, 21].

4.1.1 Classi

Considerato un generico processo decisionale di Markov, si è già osservato in sezione 2.4 che una determinata politica stazionaria μ induce nel processo una catena di Markov con matrice delle probabilità di transizione P_μ . Vale, poi, la seguente definizione.

Definizione 4.1.1 *Fissati due stati i e j di un processo decisionale di Markov, si dice che j è accessibile dallo stato i se esiste una politica stazionaria μ e un intero k tali che*

$$P[s_{t+k} = j | s_t = i, \mu] > 0, \quad t \in \mathcal{T},$$

cioè se, applicando la politica stazionaria μ , è positiva la probabilità di raggiungere lo stato j dallo stato i in un qualche numero k di passi.

I processi decisionali di Markov possono essere classificati valutando due aspetti fondamentali:

- La struttura delle catene di Markov indotte nel processo da *tutte* le politiche stazionarie.
- L'accessibilità tra gli stati del processo indotta da *alcune* politiche stazionarie

Un qualsiasi processo decisionale di Markov è detto *generale*, dopo di che può essere:

- *Ricorrente*: se la catena di Markov indotta da *ogni* politica stazionaria consiste di una sola classe ricorrente.

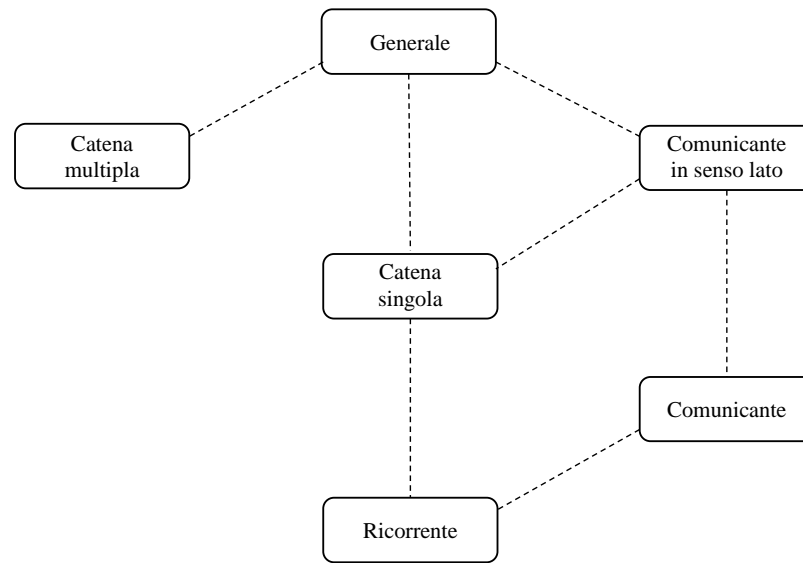


Figura 4.1 Classificazione dei processi decisionali di Markov.

- *A catena singola o unichain*: nel caso la catena di Markov indotta da *ogni* politica stazionaria consista di una sola classe ricorrente più un insieme eventualmente vuoto di stati transitori.
- *Comunicante*: se comunque considerati due stati i e j del processo, lo stato j è accessibile dallo stato i secondo la definizione 4.1.1
- *Comunicante in senso lato*: se l'insieme degli stati può essere partizionato in due sottoinsiemi \mathcal{S}_t e \mathcal{S}_c tali che: comunque fissati due stati i e j in \mathcal{S}_c , j sia accessibile da i e *tutti* gli stati in \mathcal{S}_t siano transitori per *tutte* le politiche stazionarie applicabili.
- *A catena multipla o multichain*: nel caso esista *almeno una* politica stazionaria tale da indurre una catena di Markov contenente due o più classi ricorrenti.

In figura 4.1 è rappresentata la classificazione appena introdotta e le relazioni che intercorrono tra le diverse classi. Quest'ultime sono mostrate in ordine di generalità: le classi più in alto sono quelle più generali mentre quelle più in basso sono le più specifiche. Il collegamento tra due classi indica che quella in posizione inferiore è contenuta in quella soprastante.

In base alle definizioni date, è immediato osservare che il concetto di processo decisionale a catena singola generalizza quello di processo decisionale ricorrente così come il concetto di processo comunicante in senso lato generalizza quello di processo comunicante. Inoltre, un processo decisionale ricorrente è certamente anche comunicante e un processo a catena singola è sicuramente comunicante in senso lato. Infine, è importante notare che un processo decisionale a catena multipla può essere o meno comunicante o comunicante in senso lato mentre non è certamente ricorrente o a catena singola.

4.1.2 Metodo di classificazione

Dato un processo decisionale di Markov, si presenta il problema di determinare a quale classe esso appartenga. A tal fine è molto utile il noto *algoritmo di Fox-Landi* [8, 21] applicabile ad una qualsiasi catena di Markov a stati finiti con spazio degli stati \mathcal{S} di cardinalità $|\mathcal{S}|$. Tale algoritmo consente di determinare le classi ricorrenti e gli stati transitori della catena eseguendo un numero $O(|\mathcal{S}|^2)$ di confronti.

Tipicamente il problema più complesso è determinare se il processo decisionale considerato sia ricorrente o a catena singola, dato che è necessario valutare le catene di Markov indotte da tutte le politiche stazionarie. Per problemi di piccole dimensioni possono essere analizzate manualmente tutte le politiche stazionarie, mentre per processi più complessi può essere applicato l'algoritmo di Fox-Landi alla catena di Markov corrispondente ad ogni politica stazionaria. Tuttavia il numero di politiche stazionarie è $\prod_{i \in \mathcal{S}} |\mathcal{A}(i)| \leq M^{|\mathcal{S}|}$, dove $M = \max_{i \in \mathcal{S}} |\mathcal{A}(i)|$; dunque il procedimento richiederebbe nel caso pessimo $O(M^{|\mathcal{S}|} |\mathcal{S}|^2)$ confronti che, nel caso di molte applicazioni, rappresentano una complessità troppo elevata. Molto spesso, quindi, si rinuncia a stabilire se il processo sia ricorrente o a catena singola evitando di utilizzare tecniche di ottimizzazione che sfruttino tali proprietà. In alcune situazioni, però, il problema viene risolto dimostrando che il processo decisionale è a catena multipla, dato che in tal caso è sufficiente individuare una sola politica stazionaria opportuna; in questo modo si è sicuri che il processo non possa essere né ricorrente né a catena singola.

Stabilire se un processo decisionale di Markov sia comunicante o comunicante in

senso lato è, invece, molto più semplice. Un possibile approccio da seguire è presentato nella seguente proposizione [21].

Proposizione 4.1.1 *Per determinare se un generico processo decisionale di Markov è comunicante, comunicante in senso lato o né l'uno né l'altro è possibile seguire il seguente procedimento:*

1. *Si definisca la matrice Y nel seguente modo: $[Y]_{ij} = y(j|i) = 1$ se $p(j|i, a) > 0$ per qualche azione $a \in \mathcal{A}(i)$, altrimenti $y(j|i) = 0$.*
2. *Si applichi l'algoritmo di Fox-Landi alla matrice Y .*
3. *Se l'algoritmo individua una sola classe ricorrente e nessun stato transitorio, si classifichi il processo decisionale come comunicante; se viene individuata una sola classe ricorrente più alcuni stati transitori, si proceda col punto 4; altrimenti il processo decisionale non è né comunicante né comunicante in senso lato.*
4. *Si ponga $c(i) = 1$ se lo stato i è nella classe ricorrente individuata, altrimenti $c(i) = 0$. Si ripeta la seguente operazione fino a quando $c(i)$ non cambia per nessun $i \in \mathcal{S}$: per ogni $i \in \mathcal{S}$ tale che $c(i) = 0$, si ponga $c(i) = 1$ se $\sum_{j \in \mathcal{S}} p(j|i, a)c(j) > 0$ per tutte le $a \in \mathcal{A}(i)$.*
5. *Se $c(i) = 1$ per tutti gli stati $i \in \mathcal{S}$, si classifichi il processo decisionale come comunicante in senso lato, altrimenti non è né comunicante né comunicante in senso lato.*

4.2 Processi comunicanti in senso lato e costo medio per slot

Si consideri un processo decisionale di Markov comunicante in senso lato. Per definizione l'insieme degli stati \mathcal{S} può essere partizionato in due sottoinsiemi \mathcal{S}_t e \mathcal{S}_c in modo tale che tutti gli stati in \mathcal{S}_t siano transitori per tutte le politiche stazionarie e che per ogni coppia di stati i e j in \mathcal{S}_c lo stato j sia accessibile da i .

Per un processo di questo tipo vale il seguente risultato di fondamentale importanza per l'ottimizzazione secondo il criterio del costo medio per slot.

Teorema 4.2.1 *Dato un processo decisionale di Markov comunicante in senso lato, il costo medio per slot ottimo è costante, ovvero è lo stesso per tutti i possibili stati iniziali.*

Dimostrazione. Sia μ una politica stazionaria ottima secondo il criterio del costo medio per slot la cui esistenza è assicurata dalla proposizione 2.4.8; allora $J_\mu(i)$ con $i \in \mathcal{S}$ è il costo medio ottimo a partire dallo stato iniziale i .

Si dimostra prima di tutto che il costo medio per slot ottimo è lo stesso a partire da tutti gli stati in \mathcal{S}_c . Si assuma per assurdo il contrario, cioè che entrambi gli insiemi

$$M = \left\{ i \in \mathcal{S}_c \mid J_\mu(i) = \max_{j \in \mathcal{S}} \{J_\mu(j)\} \right\}$$

ed il suo complementare in \mathcal{S}_c , $\bar{M} = \{i \in \mathcal{S}_c \mid i \notin M\}$, siano non vuoti. Per definizione di processo comunicante in senso lato, comunque fissati due stati $i \in M$ e $j \in \bar{M}$, esiste una politica stazionaria μ' e un intero k tali che

$$P[s_{t+k} = j \mid s_t = i, \mu'] > 0, \quad t \in \mathcal{T}.$$

Ciò significa che nella catena di Markov indotta nel processo dalla politica stazionaria μ' esiste una transizione con probabilità positiva da uno stato $m \in M$ ad uno stato $\bar{m} \in \bar{M}$; quindi esistono $m \in M$ e $\bar{m} \in \bar{M}$ tali che

$$[P_{\mu'}]_{m\bar{m}} = P[s_{t+1} = \bar{m} \mid s_t = m, \mu'] > 0, \quad t \in \mathcal{T}.$$

Ma allora segue che l' m -esima componente di $P_{\mu'} J_\mu$ è strettamente minore dell' m -esima componente di J_μ che è $\max_{j \in \mathcal{S}} \{J_\mu(j)\}$. Ma questo contraddice l'equazione di ottimalità (4.1) secondo la quale, dato che μ è ottima, vale

$$J_\mu = P_\mu J_\mu \leq P_{\mu'} J_\mu.$$

Dunque l'ipotesi posta per assurdo non è valida e quindi il costo medio per slot ottimo è il medesimo a partire da tutti gli stati in \mathcal{S}_c .

Inoltre, dato che per definizione di processo comunicante in senso lato tutti gli stati in \mathcal{S}_t sono transitori per tutte le politiche stazionarie, lo sono anche rispetto alla politica stazionaria ottima μ . Quindi se lo stato di partenza è un qualsiasi stato in \mathcal{S}_t e viene

applicata la politica μ , il sistema raggiunge uno stato di \mathcal{S}_c in un numero medio di transizioni finito. Ma allora il costo medio per slot di μ a partire da un qualsiasi stato in \mathcal{S}_t è uguale al costo medio per slot comune a tutti gli stati in \mathcal{S}_c . ■

4.2.1 Proprietà del modello proposto per il sensore radio

Alla luce della classificazione introdotta in sezione 4.1 è interessante capire, a questo punto, quali siano le proprietà e a quale classe appartenga il processo decisionale di Markov proposto nel capitolo 3 per rappresentare il sensore radio.

Si può facilmente verificare che il processo decisionale è a *catena multipla*. Si consideri infatti la politica stazionaria che prevede che in qualsiasi stato in cui si trovi il dispositivo in un determinato slot non venga eseguita nessuna operazione. Nel caso di radio e sensore indipendenti ciò significa eseguire in ogni stato l'azione (*OFF, ID*), mentre con radio e sensore mutuamente esclusivi l'azione scelta è sempre (*ID*). In questo caso è immediato osservare che durante l'evoluzione del sistema la situazione dei buffer rimane inalterata rispetto allo stato di partenza. La catena di Markov indotta dalla politica stazionaria considerata è allora costituita da tante classi ricorrenti quante sono le possibili combinazioni di valori per le variabili di stato w e r (almeno quattro). Dunque esiste una politica stazionaria la cui corrispondente catena di Markov è composta da più di una classe ricorrente, quindi per definizione il processo è a catena multipla. Ciò assicura, conseguentemente, che il processo non sia né ricorrente né a catena singola.

Tuttavia il risultato fondamentale che interessa ai fini dell'ottimizzazione è che il processo decisionale per il sensore radio è *comunicante in senso lato*; tale proprietà può essere verificata adottando la procedura illustrata nella proposizione 4.1.1. Il fatto che il modello proposto sia di questo tipo consente di sfruttare il risultato presentato nel teorema 4.2.1 che assicura che il costo medio per slot ottimo sia costante.

Si è già visto nella proposizione 2.4.10 che, nel caso in cui il costo medio per slot ottimo sia lo stesso per qualsiasi stato iniziale, le equazioni di ottimalità (4.1) e (4.2) si riducono alla singola equazione

$$\lambda + h(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) h(j) \right], \quad i \in \mathcal{S}. \quad (4.3)$$

Una volta determinati uno scalare λ e un vettore h che soddisfino (4.3), si conosce il costo medio per slot ottimo e costante, pari a λ , e la corrispondente politica stazionaria ottima μ^* che è quella che realizza il minimo in (4.3).

In base a quanto detto, il calcolo delle politiche ottime per il sensore radio può essere eseguito individuando un algoritmo di ottimizzazione che consenta di risolvere l'equazione (4.3).

4.3 Procedura di ottimizzazione

Una volta stabilito che il processo decisionale di Markov proposto nel capitolo 3 è comunicante in senso lato e che l'equazione di ottimalità da considerare per individuare la politica ottima secondo il criterio del costo medio per slot è la (4.3), l'obiettivo a questo punto è quello di stabilire una procedura da seguire per la risoluzione dell'equazione (4.3).

Con questo proposito è utile introdurre il concetto di *costo totale atteso* per un processo decisionale di Markov ad orizzonte finito, a cui si è già accennato in sezione 2.4.3. Si consideri, quindi, un processo ad orizzonte finito N con spazio degli stati \mathcal{S} , costi sul singolo slot g e costi finali nulli, dove i costi finali sono quelli accumulati nell'istante finale N in base allo stato occupato. Fissata una generica politica ammissibile $\pi = \{\mu_0, \mu_1, \dots, \mu_{N-1}\}$ per tale processo, si definisce costo totale atteso della politica π a partire dallo stato $i \in \mathcal{S}$:

$$V_{\pi, N}(i) = E \left[\sum_{t=0}^{N-1} g(s_t, \mu_t(s_t)) \middle| s_0 = i \right], \quad i \in \mathcal{S}. \quad (4.4)$$

Il costo totale atteso *ottimo* è il minimo valutato tra tutte le politiche ammissibili per il processo considerato:

$$V_N(i) = \min_{\pi \in \Pi} V_{\pi, N}(i), \quad i \in \mathcal{S}. \quad (4.5)$$

Si può dimostrare [25] che il costo totale atteso ottimo per un processo ad orizzonte finito N soddisfa l'equazione di ottimalità

$$V_N(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) V_{N-1}(j) \right], \quad i \in \mathcal{S}, N \geq 1. \quad (4.6)$$

Proprio dall'equazione (4.6) nasce l'idea per una procedura di ottimizzazione applicabile ad un processo decisionale ad orizzonte infinito per il calcolo della politica ottima secondo il criterio del costo medio per slot [25]. Si torni a considerare, quindi, un processo decisionale ad orizzonte infinito come quello proposto nel capitolo 3; si supponga di indicare con x uno specifico stato di \mathcal{S} e per ogni stato $i \in \mathcal{S}$ si definisca $r_N(i) = V_N(i) - V_N(x)$. Sottraendo ad entrambi i membri di (4.6) la quantità $V_{N-1}(x)$ e aggiungendo e togliendo $V_N(x)$ al membro sinistro, l'equazione (4.6) può essere riscritta come

$$[V_N(x) - V_{N-1}(x)] + r_N(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) r_{N-1}(j) \right], \quad i \in \mathcal{S}, N \geq 1. \quad (4.7)$$

Si assuma, ora, che sia possibile dimostrare che il termine tra parentesi nella parte sinistra dell'equazione (4.7) tenda ad uno scalare J per $N \rightarrow \infty$. Si ipotizzi, inoltre, che r_N converga per $N \rightarrow \infty$ ad un certo vettore r . Allora, eseguendo il limite per $N \rightarrow \infty$ in entrambi i membri di (4.7) si otterrebbero uno scalare $\lambda = J$ e un vettore $h = r$ tali da soddisfare l'equazione di ottimalità (4.3) che in questo modo sarebbe quindi risolta.

Affinché questa idea generale possa essere tradotta in un algoritmo di ottimizzazione effettivamente valido è necessario stabilire quali ipotesi debbano essere soddisfatte per garantire la convergenza di r_N e $[V_N(x) - V_{N-1}(x)]$. A tal fine è importante il seguente lemma.

Lemma 4.3.1 *Si considerino uno scalare J e un vettore h tali da soddisfare l'equazione di ottimalità (4.3); lo scalare J è quindi il costo medio per slot ottimo e costante. Si definisca poi*

$$d_N(i) = h(i) + NJ - V_N(i), \quad i \in \mathcal{S}, N \geq 0. \quad (4.8)$$

Assumendo che μ sia una politica stazionaria ottima tale da indurre nel processo una catena di Markov con una classe ricorrente aperiodica R , allora esiste una costante D finita tale che $\lim_{N \rightarrow \infty} d_N(i) = D$ per $i \in R$.

Dimostrazione. Si dimostra innanzitutto che

$$\sum_{j \in \mathcal{S}} P_{ij}(\mu(i)) d_{N-1}(j) \leq d_N(i), \quad i \in R, N \geq 1. \quad (4.9)$$

Dall'equazione (4.6) segue immediatamente che

$$V_N(i) \leq g(i, \mu(i)) + \sum_{j \in \mathcal{S}} p_{ij}(\mu(i)) V_{N-1}(j), \quad i \in \mathcal{S}, N \geq 1. \quad (4.10)$$

A partire dall'equazione (4.3) si ottiene invece

$$J + h(i) \leq g(i, \mu(i)) + \sum_{j \in \mathcal{S}} p_{ij}(\mu(i)) h(j), \quad i \in \mathcal{S}. \quad (4.11)$$

Poiché la politica μ è ottima, è facile verificare [25] che la disuguaglianza (4.11) vale come uguaglianza per $i \in R$, cioè $J + h(i) = g(i, \mu(i)) + \sum_{j \in \mathcal{S}} p_{ij}(\mu(i)) h(j)$ per $i \in R$. Sottraendo a quest'ultima uguaglianza la disuguaglianza (4.10) si ricava

$$J + h(i) - V_N(i) \geq \sum_{j \in \mathcal{S}} p_{ij}(\mu(i)) [h(j) - V_{N-1}(j)], \quad i \in R, N \geq 1. \quad (4.12)$$

Aggiungendo la quantità $(N-1)J$ ad entrambi i membri di (4.12) si ottiene direttamente la (4.9).

Nel prosieguo della dimostrazione sarà omesso dalla notazione l'esplicito riferimento alla politica μ : si indica semplicemente con p_{ij} la probabilità di transizione in un passo da i a j data la politica μ e, allo stesso modo, $p_{ij}^{(M)}$ rappresenta la probabilità di transizione in M passi data μ .

Si può dimostrare [25] che la funzione d_N è limitata: esiste una funzione $L(i)$ non negativa e finita, definita su \mathcal{S} , tale che la sequenza di funzioni $d_N(i)$ su \mathcal{S} soddisfi la relazione $-L(i) \leq d_N(i) \leq L(i)$, $N \geq 0$, $i \in \mathcal{S}$. Ma allora esistono una sottosequenza N_k e una funzione $d(i)$, con $-L(i) \leq d(i) \leq L(i)$, tali che $\lim_{k \rightarrow \infty} d_{N_k}(i) = d(i)$, $i \in \mathcal{S}$.

A questo punto, fissando arbitrariamente N ed iterando M volte la relazione (4.9) si ottiene

$$\sum_{j \in \mathcal{S}} p_{ij}^{(M)} d_N(j) \leq d_{N+M}(i), \quad i \in R, N \geq 0, M \geq 1. \quad (4.13)$$

Mantenendo fisso N si esegue il limite per $M \rightarrow \infty$ attraverso valori tali che $N+M$ appartengano alla sottosequenza N_k ; tenendo conto della aperiodicità della classe R e

del fatto che si tratta di una classe ricorrente, si ricava

$$\sum_{j \in R} \bar{p}_j d_N(j) \leq d(i), \quad i \in R, N \geq 0, \quad (4.14)$$

dove \bar{p}_j è la probabilità a lungo termine che il sistema occupi lo stato j . Passando al limite per $N = N_k \rightarrow \infty$ si ottiene

$$\sum_{j \in R} \bar{p}_j d(j) \leq d(i), \quad i \in R. \quad (4.15)$$

La disuguaglianza (4.15) evidenzia che ciascun valore della funzione $d(i)$, $i \in R$, è maggiore o uguale ad una combinazione convessa dei suoi valori in R ; questo è possibile se e solo se $d(i)$ è una costante in R , cioè se e solo se $d(i) = D$, $i \in R$.

Si consideri, infine, una diversa sottosequenza N_u tale che $\lim_{u \rightarrow \infty} d_{N_u}(i) = c(i)$, con $c(i)$ funzione costante e pari a C in R . Allora eseguendo il limite per $N = N_u \rightarrow \infty$ in (4.14) si ricava

$$C = \sum_{j \in R} \bar{p}_j c(j) \leq D.$$

Dato che quest'ultimo passo può essere ripetuto in ordine inverso portando alla relazione $D \leq C$, non può che essere $D = C$. Ma allora si ha effettivamente $\lim_{N \rightarrow \infty} d_N(i) = D$ per $i \in R$. ■

Alla luce del risultato dimostrato nel lemma 4.3.1 si introduce, ora, l'assunzione fondamentale su cui si basa la validità dell'algoritmo di ottimizzazione che sarà successivamente presentato.

Assunzione OPA *Se μ è una politica stazionaria ottima, allora ogni classe ricorrente della catena di Markov indotta da μ è aperiodica.*

L'appellativo *OPA (Optimal Policies are Aperiodic)* è utilizzato per sottolineare che, in caso questa assunzione sia valida, ogni politica stazionaria ottima induce esclusivamente classi ricorrenti aperiodiche.

L'assunzione OPA è di cruciale importanza per il seguente teorema.

Teorema 4.3.2 *Si consideri un processo decisionale di Markov tale che il costo medio per slot ottimo sia costante e pari a J e che sia valida l'assunzione OPA; inoltre, sia x un qualsiasi stato in \mathcal{S} . Allora esistono*

$$\lim_{N \rightarrow \infty} [V_N(x) - V_{N-1}(x)] = J, \quad (4.16)$$

$$\lim_{N \rightarrow \infty} r_N(i) = r(i), \quad i \in \mathcal{S}, \quad (4.17)$$

e, quindi, l'equazione (4.7) può essere utilizzata per calcolare una soluzione dell'equazione di ottimalità (4.3).

Dimostrazione. Sia μ una politica stazionaria ottima secondo il criterio del costo medio per slot la cui esistenza è assicurata dalla proposizione 2.4.8; per quanto detto in sezione 2.4.1, la catena di Markov indotta da μ , essendo una catena a stati finiti, contiene almeno una classe ricorrente. Si indichi con R una delle classi ricorrenti e si supponga per il momento $x \in R$; nella parte finale della dimostrazione si vedrà che x può essere scelto arbitrariamente in \mathcal{S} .

Le relazioni

$$r_N(i) = h(i) - h(x) - d_N(i) + d_N(x), \quad (4.18)$$

$$V_N(x) - V_{N-1}(x) = J - d_N(x) + d_{N-1}(x), \quad (4.19)$$

consentono di tradurre i risultati relativi a d_N in risultati validi per le quantità presenti nell'equazione (4.7).

Dato che $x \in R$ e che la classe ricorrente R è aperiodica per l'assunzione OPA, per il lemma 4.3.1 esiste una costante D finita tale che $\lim_{N \rightarrow \infty} d_N(x) = D$. Allora passando al limite per $N \rightarrow \infty$ nell'equazione (4.19) si ottiene direttamente $\lim_{N \rightarrow \infty} [V_N(x) - V_{N-1}(x)] = J$.

Ricordando che la funzione d_N su \mathcal{S} è limitata [25], dall'equazione (4.18) segue che anche la funzione r_N su \mathcal{S} è limitata. Siano allora r_* e r^* rispettivamente il limite inferiore e il limite superiore di r_N .

Eseguendo, a questo punto, il limite inferiore in entrambi i membri dell'equazione (4.7) si ottiene

$$J + r_*(i) \geq \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) r_*(j) \right], \quad i \in \mathcal{S}. \quad (4.20)$$

Indicata con μ' la politica stazionaria che realizza il minimo nella parte destra di equazione (4.20), è possibile dimostrare [25] che μ' è una politica stazionaria ottima.

Ritornando ora all'equazione (4.7) segue che

$$[V_N(x) - V_{N-1}(x)] + r_N(i) \leq g(i, \mu'(i)) + \sum_{j \in \mathcal{S}} p_{ij}(\mu'(i)) r_{N-1}(j), \quad i \in \mathcal{S}, N \geq 1. \quad (4.21)$$

Passando al limite superiore in entrambi i membri di (4.21) si ricava

$$J + r^*(i) \leq g(i, \mu'(i)) + \sum_{j \in \mathcal{S}} p_{ij}(\mu'(i)) r^*(j), \quad i \in \mathcal{S}. \quad (4.22)$$

Utilizzando la relazione (4.22) e tenendo conto che μ' realizza il minimo in (4.20) si ottiene

$$g(i, \mu'(i)) - J + \sum_{j \in \mathcal{S}} p_{ij}(\mu'(i)) r_*(j) \leq r_*(i) \leq r^*(i) \leq g(i, \mu'(i)) - J + \sum_{j \in \mathcal{S}} p_{ij}(\mu'(i)) r^*(j).$$

Posto $\omega = r^* - r_* \geq 0$ segue che

$$\omega(i) \leq \sum_{j \in \mathcal{S}} p_{ij}(\mu'(i)) \omega(j), \quad i \in \mathcal{S}. \quad (4.23)$$

Siano, ora, U_1, U_2, \dots, U_M le classi ricorrenti aperiodiche indotte dalla politica μ' e sia $q_m(i)$ la probabilità di raggiungere la classe U_m , $m = 1, 2, \dots, M$, a partire dallo stato i data μ' . Iterando la relazione (4.23) e passando al limite si ottiene facilmente

$$\omega(i) \leq \sum_{m=1}^M q_m(i) \left(\sum_{j \in U_m} \bar{p}_j \omega(j) \right), \quad i \in \mathcal{S}, \quad (4.24)$$

dove \bar{p}_j è la probabilità a lungo termine che il sistema occupi lo stato $j \in U_m$ a partire da un generico stato in U_m data μ' .

Concentrando l'attenzione su una particolare classe U_m , essa è aperiodica per l'assunzione OPA e dunque per il lemma 4.3.1 esiste una costante E tale che $\lim_{N \rightarrow \infty} d_N(j) = E$ per $j \in U_m$. Poiché $\lim_{N \rightarrow \infty} d_N(x) = D$ segue da (4.18) che esiste $\lim_{N \rightarrow \infty} r_N(j) = h(j) - h(x) - E + D$, $j \in U_m$; visto che il limite esiste si può concludere che $s(j) = 0$, $j \in U_m$.

Dato che lo stesso ragionamento può essere fatto per ogni classe U_m , con $m = 0, 1, \dots, M$, il termine destro della disuguaglianza (4.24) è necessariamente nullo, cioè

$\omega(i) \leq 0$. D'altra parte per definizione $\omega(i) \geq 0$ e quindi deve essere $\omega(i) = 0$ per ciascun $i \in \mathcal{S}$. Ma questo assicura che $r(i) = \lim_{N \rightarrow \infty} r_N(i)$ esista per ogni $i \in \mathcal{S}$ e la prova è conclusa nel caso lo stato x sia uno stato di una classe ricorrente indotta da μ .

Si assuma, invece, che lo stato di riferimento scelto sia un arbitrario $y \in \mathcal{S}$. Si può scrivere allora

$$\begin{aligned} V_N(y) - V_{N-1}(y) &= [V_N(y) - V_N(x)] + [V_N(x) - V_{N-1}(x)] - [V_{N-1}(y) - V_{N-1}(x)] \\ &= r_N(y) + [V_N(x) - V_{N-1}(x)] - r_{N-1}(y). \end{aligned}$$

In base a quanto dimostrato precedentemente si ottiene quindi $\lim_{N \rightarrow \infty} [V_N(y) - V_{N-1}(y)] = r(y) + J - r(y) = J$.

In modo analogo, si ha

$$\begin{aligned} V_N(i) - V_N(y) &= [V_N(i) - V_N(x)] - [V_N(y) - V_N(x)] \\ &= r_N(i) - r_N(y). \end{aligned}$$

Ma allora esiste $\lim_{N \rightarrow \infty} [V_N(i) - V_N(y)] = r(i) - r(y)$. ■

Il teorema 4.3.2 specifica in modo chiaro le condizioni che devono essere soddisfatte affinché le quantità r_N e $[V_N(x) - V_{N-1}(x)]$ in (4.7) convergano per $N \rightarrow \infty$: il costo medio per slot ottimo deve essere costante e deve essere valida la condizione OPA. In tal caso l'equazione (4.7) può essere sfruttata per risolvere l'equazione di ottimalità (4.3) mediante il seguente algoritmo di ottimizzazione, noto come *value iteration* [25] o *relative value iteration* [2].

Algoritmo di ottimizzazione *Si consideri un processo decisionale di Markov tale che il costo medio per slot ottimo sia costante e che sia valida l'assunzione OPA; inoltre, sia x un qualsiasi stato di riferimento in \mathcal{S} e sia ϵ un numero positivo arbitrariamente piccolo. Il costo medio per slot ottimo λ e la corrispondente politica stazionaria ottima μ possono essere così determinati:*

1. Si ponga $N = 0$ e $u_0 = 0$.

2. Si calcoli

$$w_N(i) = \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) u_N(j) \right], \quad i \in \mathcal{S}.$$

3. Se $N = 0$, si ponga $\delta = 1$; altrimenti $\delta = \max_{i \in \mathcal{S}} |w_N(i) - w_{N-1}(i)|$. Se $\delta < \epsilon$ si vada al passo 6.
4. Si calcoli $u_{N+1}(i) = w_N(i) - w_N(x)$, $i \in \mathcal{S}$.
5. Si torni al passo 2 sostituendo N con $N + 1$.
6. Si restituisca $\lambda = w_N(x)$ e la politica μ che realizza

$$\min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) u_N(j) \right], \quad i \in \mathcal{S}.$$

Dimostrazione. Si dimostra per induzione su N che

$$w_N(i) = r_{N+1}(i) + V_{N+1}(x) - V_N(x), \quad i \in \mathcal{S}, N \geq 0. \quad (4.25)$$

Per $N = 0$ si ha

$$w_0(i) = \min_{a \in \mathcal{A}(i)} [g(i, a)],$$

inoltre

$$\begin{aligned} r_1(i) + V_1(x) - V_0(x) &= V_1(i) - V_1(x) + V_1(x) \\ &= \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) V_0(j) \right] \\ &= \min_{a \in \mathcal{A}(i)} [g(i, a)], \end{aligned}$$

quindi la relazione (4.25) è valida per $N = 0$. Si assume, ora, che la (4.25) sia soddisfatta per N e se ne dimostra la validità per $N + 1$:

$$\begin{aligned} w_{N+1}(i) &= \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) u_{N+1}(j) \right] \\ &= \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) (w_N(j) - w_N(x)) \right], \end{aligned}$$

dove

$$\begin{aligned} w_N(j) - w_N(x) &= r_{N+1}(j) + V_{N+1}(x) - V_N(x) - r_{N+1}(x) - V_{N+1}(x) + V_N(x) \\ &= r_{N+1}(j). \end{aligned}$$

Ma allora

$$\begin{aligned} w_{N+1}(i) &= \min_{a \in \mathcal{A}(i)} \left[g(i, a) + \sum_{j \in \mathcal{S}} p_{ij}(a) r_{N+1}(j) \right] \\ &= r_{N+2}(i) + V_{N+2}(x) - V_{N+1}(x). \end{aligned}$$

La relazione (4.25) è quindi valida per ogni $N \geq 0$ e in particolare si ha $w_N(x) = V_{N+1}(x) - V_N(x)$.

Si può poi dimostrare che

$$u_N(i) = r_N(i), \quad i \in \mathcal{S}, N \geq 0. \quad (4.26)$$

Per $N = 0$, $u_0(i) = 0$ e $r_0(i) = V_0(i) - V_0(x) = 0$ e la (4.26) è valida. Considerato ora $N \geq 1$ si ha

$$\begin{aligned} u_N(i) &= w_{N-1}(i) - w_{N-1}(x) \\ &= r_N(i) + V_N(x) - V_{N-1}(x) - r_N(x) - V_N(x) + V_{N-1}(x) \\ &= r_N(i), \end{aligned}$$

quindi la relazione (4.26) è soddisfatta per ogni $N \geq 0$.

Alla luce di quanto appena dimostrato, la validità dell'algoritmo è conseguenza del teorema 4.3.2. L'algoritmo calcola iterativamente le quantità $w_N(x) = V_{N+1}(x) - V_N(x)$ e $u_N = r_N$ fino a quando convergono rispettivamente a J e a r . Una volta avvenuta la convergenza, garantita dal teorema 4.3.2, si hanno a disposizione uno scalare $\lambda = J$ e un vettore $h = r$ che soddisfano l'equazione di ottimalità (4.3). Dunque λ è il costo medio per slot ottimo e la politica ottima è quella che realizza il minimo. ■

L'algoritmo appena proposto è un valido strumento di ottimizzazione per processi decisionali di Markov con costo medio per slot ottimo costante e tali da soddisfare l'assunzione OPA.

Si è già visto in sezione 4.2.1 che il processo decisionale proposto nel capitolo 3 per rappresentare il sensore radio è comunicante in senso lato; in base al teorema 4.2.1, questo è sufficiente per garantire che il costo medio per slot ottimo sia costante.

D'altra parte, per come è stato definito il processo decisionale, in particolare le probabilità di transizione e le azioni ammissibili, non sono presenti i presupposti

affinché una politica stazionaria possa indurre una catena di Markov con classi ricorrenti periodiche. Pertanto, anche l'assunzione OPA è soddisfatta.

L'algoritmo presentato può quindi essere utilizzato per calcolare le politiche ottime per il sensore radio al variare dei parametri che caratterizzano il modello.

5

Simulazioni, analisi e risultati

Nei capitoli precedenti, dopo aver proposto un modello per la rappresentazione del funzionamento di un sensore radio alimentato da una fonte di energia rinnovabile, si è individuato un algoritmo, ad esso applicabile, che consenta di determinare le politiche di scheduling ottime per il dispositivo al variare delle condizioni in cui si trova ad operare. In questo capitolo si concentra l'attenzione sullo studio degli effetti che le variazioni di alcuni parametri del modello hanno sulle politiche ottime calcolate mediante l'algoritmo di ottimizzazione. Nella prima parte del capitolo si presentano alcuni risultati ottenuti mediante simulazioni: al variare di alcuni parametri si calcolano le corrispondenti politiche ottime e si simula l'evoluzione del dispositivo in accordo con tali politiche valutando quali siano le conseguenze a livello di throughput. Nella seconda parte, si analizzano direttamente le politiche ottime evidenziandone alcune caratteristiche particolarmente interessanti e valutando come queste siano influenzate dalla scelta dei parametri.

5.1 Parametri del modello

Nel capitolo 3 si è analizzato in dettaglio il modello proposto in questa tesi per rappresentare un sensore radio alimentato da una fonte di energia rinnovabile. Per come è stato definito il modello, il problema di determinare quale sia la politica di scheduling ottima che il dispositivo debba seguire per massimizzare le sue prestazioni si è tradotto nel problema di individuare la politica ottima secondo il criterio del costo medio per slot del processo decisionale di Markov rappresentativo del sistema. Nel capitolo 4 si è poi sottolineato come il modello proposto soddisfi le ipotesi che assicurano l'applicabilità dell'algoritmo di ottimizzazione illustrato in sezione 4.3. Tale algoritmo, una volta fissati i parametri del sistema, può dunque essere utilizzato per determinare la corrispondente politica ottima secondo il criterio del costo medio per slot.

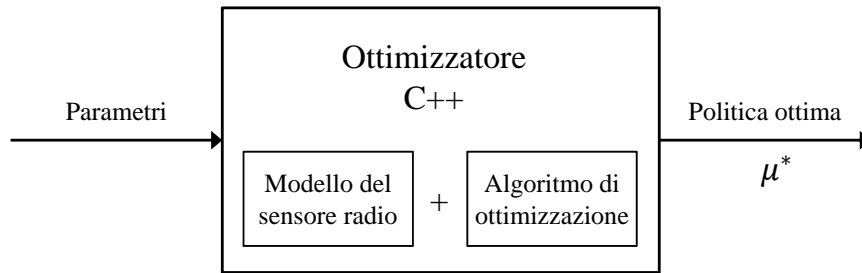


Figura 5.1 Ingresso e uscita dell'ottimizzatore realizzato in C++.

Alla luce di questo, sono stati implementati in C++ sia il processo decisionale rappresentativo del sensore radio che il relativo algoritmo di ottimizzazione. In questo modo è stato quindi realizzato un *ottimizzatore*, cioè uno strumento di calcolo in grado di determinare le politiche ottime. Come mostrato in figura 5.1, all'ottimizzatore vengono forniti in ingresso i parametri che definiscono le condizioni operative in cui si trova il dispositivo e in uscita viene restituita la politica di scheduling ottima, in tali condizioni, risultante dall'applicazione dell'algoritmo di ottimizzazione.

I risultati presentati in questo capitolo sono ottenuti utilizzando più volte l'ottimizzatore con dati in ingresso diversi. I parametri in gioco sono quelli che caratterizzano il modello di capitolo 3; di seguito vengono riassunti specificando, per ciascuno di essi, se si tratta di un parametro il cui valore è prefissato oppure variabile in un certo intervallo.

5.1.1 Consumi energetici

In sezione 3.6 si è visto che qualsiasi attività svolta in un generico slot dal sensore radio comporta il consumo di un certo numero di unità di carica contenute nella batteria: E_{tx} è il numero di unità richiesto in uno slot per la trasmissione, E_{rx} quello per la ricezione, E_{sx} per l'attività di sensing e E_{id} per la radio in idle.

Si assume che E_{tx} sia il consumo energetico di riferimento ponendo poi:

- $E_{rx} = K_{rx}E_{tx}$,
- $E_{sx} = K_{sx}E_{tx}$,
- $E_{id} = K_{id}E_{tx}$.

In questo modo i parametri K_{rx} , K_{sx} e K_{id} consentono di esprimere facilmente la quantità di energia consumata dalle corrispondenti attività in relazione a quella spesa per la trasmissione. Una volta stabilito il valore di tali parametri il rapporto che intercorre tra E_{tx} , E_{rx} , E_{sx} e E_{id} è fissato indipendentemente dagli effettivi valori assunti da quest'ultimi.

I chip radio tipicamente utilizzati nei dispositivi (come ad esempio il Chipcon CC2420 [29]) portano ad un dispendio di energia analogo in fase di trasmissione e ricezione. Per questo motivo è sensato fissare $K_{rx} = 1$ e assumere, quindi, che l'operazione di ricezione richieda lo stesso numero di unità di carica per slot speso per la trasmissione.

Mantenere la radio in idle, invece, comporta un dispendio di energia del tutto trascurabile rispetto a quello dovuto alla trasmissione. Pertanto è lecito porre $K_{id} = 0$, considerando nullo il consumo energetico per slot E_{id} della radio in idle.

Infine, il numero di unità di carica E_{sx} richieste in uno slot per l'attività di sensing dipende fortemente sia dal tipo di rilevazione effettuata che dal particolare sensore utilizzato per compierla. Pertanto, contrariamente a K_{rx} e K_{id} , il valore del parametro K_{sx} non può essere genericamente stabilito ma deve essere fissato in base allo specifico dispositivo considerato e alla particolare attività di monitoraggio eseguita. Ciò che è possibile affermare in generale, in base a quanto si può vedere nelle applicazioni reali, è che l'energia necessaria per l'attività di sensing può variare da valori molto inferiori a quelli relativi all'energia spesa per la trasmissione fino a valori comparabili o anche superiori. Conseguentemente è ragionevole assumere che i valori che il parametro K_{sx} può assumere siano quelli compresi nell'intervallo $(0, 1.5]$.

5.1.2 Batteria

La capacità N della batteria, ovvero il massimo numero di unità di carica che essa contiene nello stato di completa carica, può essere anch'essa espressa in funzione di E_{tx} :

- $N = K_N E_{tx}$.

In questo caso il parametro K_N rappresenta il numero massimo di trasmissioni che la batteria è in grado di sostenere prima di esaurirsi completamente nel caso non venga ricaricata in alcun modo. L'effettivo valore di K_N dipende dalla specifica batteria che viene utilizzata dal sensore radio e dalla particolare applicazione considerata. Nei risultati che verranno proposti in questo capitolo si assume, a titolo esemplificativo, che la batteria a disposizione sia tale che $K_N = 1000$.

Il secondo parametro che riguarda la batteria è il livello di guardia $N_{lim} = \lceil \rho N \rceil$. Il valore di ρ può essere opportunamente fissato per variare la profondità massima a cui si desidera venga scaricata la batteria dall'attività del sensore radio. Chiaramente deve essere $\rho \in [0, 1]$.

5.1.3 Fonte di energia rinnovabile

In sezione 3.5 si è discusso ampiamente dei modelli proposti in questa tesi per rappresentare la fonte di energia rinnovabile.

Il primo parametro fondamentale che caratterizza la fonte di energia è E_{MAX} , che indica il massimo numero di unità di carica con il quale essa è in grado di caricare la batteria in uno slot. È conveniente esprimere anche E_{MAX} in funzione di E_{tx} :

- $E_{MAX} = K_{MAX} E_{tx}$.

In questo modo, mediante K_{MAX} , è possibile indicare il numero massimo di unità di carica con cui può essere rifornita la batteria in uno slot in relazione al numero di unità necessarie per compiere la trasmissione di un pacchetto.

Il valore di K_{MAX} , e quindi di E_{MAX} , è condizionato da molteplici fattori: tipo di sorgente di energia, meccanismo di conversione, efficienza di conversione, posizione del sensore radio. Considerata ad esempio l'energia solare, il valore di E_{MAX} dipende dalla posizione del dispositivo e quindi da quanto è esposto alla radiazione solare, dalle dimensioni del pannello solare, da quanto efficiente è la conversione dell'energia solare in energia elettrica per caricare la batteria. Per questo motivo, è chiaro che i valori effettivi di K_{MAX} ed E_{MAX} non possono essere stabiliti in generale a priori ma devono essere fissati in base alle condizioni operative contingenti in cui si trova il sensore radio considerato.

Nei risultati presentati in questo capitolo si farà riferimento ai due modelli per la fonte di energia rinnovabile più interessanti: il modello generale a stato singolo di sezione 3.5.2 e il modello a due stati di sezione 3.5.3 presentato in figura 3.5.

Il modello generale a stato singolo, in cui $X = 0$ e conseguentemente $\gamma_{00} = 1$, può essere adottato per rappresentare una qualsiasi sorgente di energia. In questo caso si suppone che in ciascuno slot il numero di unità di carica fornite dalla fonte alla batteria assuma valori nell'insieme $\Delta E_0 = \{0, 1, \dots, E_{MAX}\}$ secondo la distribuzione $q_0(\cdot)$. Quindi, per ogni slot, $q_0(e)$ indica la probabilità che la batteria venga ricaricata con e unità. Nel seguito si assumerà per la distribuzione $q_0(\cdot)$, rappresentativa della sorgente di energia, un andamento a campana sull'insieme dei valori ammissibili ΔE_0 , con ampiezza della campana modulabile mediante un parametro σ e con media m centrata nell'intervallo. Per ottenere una distribuzione di questo tipo, l'idea è quella di campionare in corrispondenza dei valori $\{0, 1, \dots, E_{MAX}\}$ la funzione densità di probabilità gaussiana (continua) con media m e deviazione standard σ :

$$f(i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(i-m)^2}{2\sigma^2}}, \quad -\infty < i < +\infty.$$

Una volta eseguito il campionamento è necessario normalizzare i valori ottenuti affinché la loro somma sia pari a 1 così che la distribuzione $q_0(\cdot)$ risultante soddisfi la condizione di normalizzazione $\sum_{e \in \Delta E_0} q_0(e) = 1$.

Per chiarezza, in figura 5.2 è mostrato l'andamento assunto da $q_0(\cdot)$, per diversi valori di σ , nel caso in cui $E_{MAX} = 8$. Come si può vedere, al decrescere di σ l'ampiezza della campana si riduce e la distribuzione risulta sempre più piccata attorno al valor medio $m = 4$, fino a giungere al caso limite di distribuzione deterministica indicata, con abuso di notazione, con $\sigma = 0$; al crescere di σ , invece, la campana si allarga progressivamente fino a giungere al caso estremo di distribuzione uniforme indicata con $\sigma = \infty$.

Adottare una distribuzione di questo tipo consente di rappresentare il grado di incertezza che si ha sul numero di unità di carica che la fonte rinnovabile fornisce alla batteria in ogni slot: valori piccoli di σ indicano poca incertezza dato che con alta probabilità il numero di unità che caricano la batteria è pari al valor medio; valori grandi di σ , viceversa, denotano maggiore incertezza poiché la distribuzione di probabilità

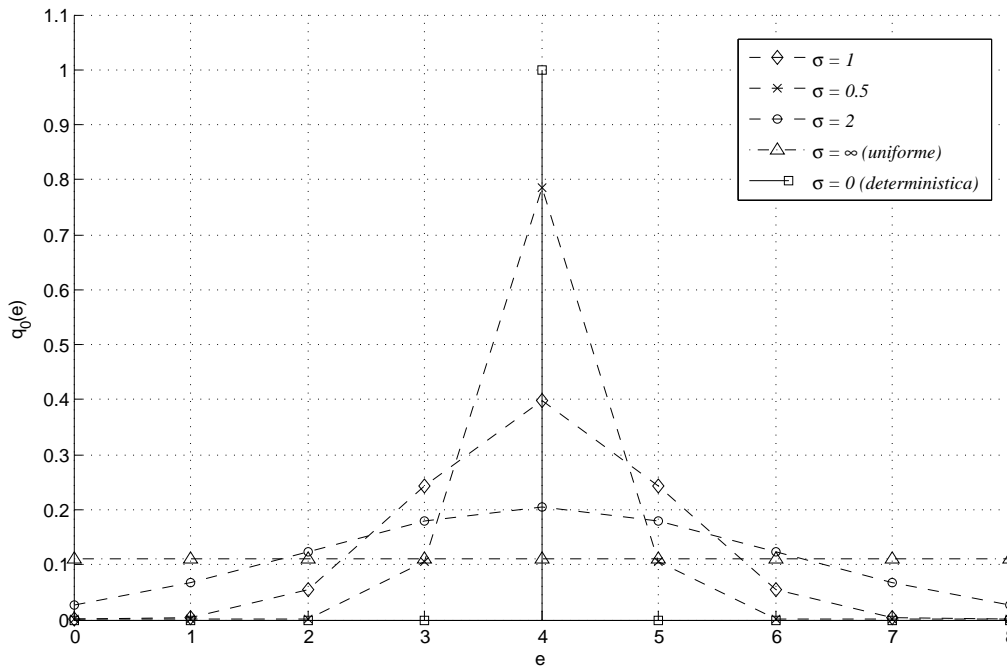


Figura 5.2 Modello a stato singolo ($X = 0$): distribuzione $q_0(\cdot)$, per diversi valori di σ , con $E_{MAX} = 8$.

è più uniforme sull'intervallo di valori ammissibili. Naturalmente nel caso limite di distribuzione deterministica l'incertezza è nulla, mentre con distribuzione uniforme l'incertezza è massima.

Il modello a due stati, con $X = 1$, è utile per rappresentare una fonte di energia che alterna fasi in cui non fornisce alcuna unità di carica alla batteria con altre in cui effettua una ricarica con flusso di energia più o meno intenso. Per questo motivo questo modello sarà adottato riferendosi esplicitamente all'energia solare, utilizzabile di giorno per alimentare la batteria del sensore radio ma assente nelle ore notturne.

In questo caso, se in un generico slot lo stato occupato dalla sorgente di energia è $x = 0$ non viene fornita nessuna unità di carica alla batteria con probabilità $q_0(0) = 1$. Se, invece, lo stato occupato è $x = 1$ la batteria viene ricaricata con un numero di unità aleatorio nell'insieme $\Delta E_1 = \{1, \dots, E_{MAX}\}$ secondo la distribuzione $q_1(\cdot)$. Nel seguito si assumerà anche per la distribuzione $q_1(\cdot)$ un andamento a campana analogo a quelli mostrati in figura 5.2, ma centrato sull'insieme di valori ammissibili $\Delta E_1 = \{1, \dots, E_{MAX}\}$. Ancora una volta una distribuzione di questo tipo consente di

rappresentare l'incertezza che si ha sul numero di unità di carica per slot che la sorgente di energia fornisce alla batteria durante le ore diurne.

Se si adotta il modello a due stati è necessario stabilire anche le probabilità di transizione con cui la fonte di energia passa dallo stato $x = 0$ allo stato $x = 1$ e viceversa. In sostanza devono essere fissati i valori di γ_{01} , γ_{10} e conseguentemente di $\gamma_{00} = 1 - \gamma_{01}$ e $\gamma_{11} = 1 - \gamma_{10}$. Tali probabilità possono essere determinate tenendo conto che il numero medio di slot durante i quali la fonte di energia rimane negli stati $x = 0$ e $x = 1$ è pari rispettivamente a $1/\gamma_{01}$ e $1/\gamma_{10}$. Facendo riferimento all'energia solare e assumendo che la durata media del dì e della notte sia $12 h$ ciascuno, si può fissare $\gamma_{01} = \gamma_{10} = 1/86400$ nel caso il singolo slot abbia durata pari a $500 ms$. Così facendo, infatti, il numero medio di slot che la sorgente trascorre nello stato $x = 0$ è pari a 86400 e quindi non fornisce unità di carica alla batteria per un tempo medio pari $86400 \times 500 ms = 43200 s = 12 h$ che è appunto la durata media della notte; analogo è il discorso per le ore diurne.

5.1.4 Buffer, traffico esogeno e locale

In ciascuno slot l'eventuale pacchetto dati proveniente dalla radio in ricezione o dall'attività di sensing svolta dal sensore viene memorizzato, rispettivamente, nel buffer per il traffico esogeno, nel primo caso, o nel buffer per il traffico locale, nel secondo.

I sensori radio comunemente utilizzati nelle applicazioni reali sono dotati di una scarsa capacità di memorizzazione dati. Per questo motivo è lecito assumere, come si farà nel resto di questo capitolo, che $W = 1$ e $R = 1$. Ciò significa supporre che sia il buffer per il traffico esogeno che quello per il traffico locale siano in grado di memorizzare al più un pacchetto ciascuno. In base a questa ipotesi, ognuno dei due buffer può essere vuoto, se non contiene nessun pacchetto, o pieno, se ne contiene uno.

Si ricorda, poi, che un altro parametro molto importante del modello su cui è possibile intervenire è $\beta \in [0, 1]$. Tale parametro rappresenta la probabilità che in un dato slot il sensore radio prediliga l'attività di rilevazione piuttosto che quella di inoltrare di traffico esterno. Variando β è possibile specificare quale delle due attività

sia più rilevante per il dispositivo e in che misura, come già spiegato ampiamente in sezione 3.6.4.

Infine, altri due parametri di cruciale importanza per il modello del sensore radio sono P_{rx} e P_{sx} . Questi due parametri specificano rispettivamente la probabilità che in un dato slot sia presente un pacchetto esogeno da ricevere e la probabilità che l'operazione di sensing compiuta in un certo slot porti all'acquisizione di un nuovo pacchetto con informazioni aggiornate. Pertanto, P_{rx} e P_{sx} modellano l'intensità del traffico esogeno e di quello locale. I valori effettivi di tali parametri dipendono da diversi fattori quali il ruolo del sensore radio nella WSN, la posizione occupata dal dispositivo nella rete, la particolare applicazione considerata.

5.2 Risultati ottenuti mediante simulazioni

Una volta fissati i valori dei parametri riassunti in sezione 5.1, l'ottimizzatore realizzato in C++ consente di determinare la politica ottima μ^* secondo il criterio del costo medio per slot. Tale politica è quella che, una volta applicata dall'unità di controllo, porta il dispositivo a massimizzare il numero medio di pacchetti trasmessi per slot nelle condizioni operative specificate dai parametri.

In questa sezione vengono presentati alcuni risultati ottenuti mediante simulazioni. Una volta stabiliti i parametri, viene prima determinata con l'ottimizzatore la politica ottima e dopo viene simulata l'evoluzione del sensore radio, in accordo col modello di capitolo 3, assumendo che le azioni compiute dal dispositivo siano quelle stabilite dalla politica ottima calcolata. Pertanto, se all'inizio di un generico slot temporale lo stato occupato dal sistema è $i \in \mathcal{S}$, l'azione che viene eseguita dal sensore radio è $\mu^*(i)$. A simulazione conclusa si valutano le prestazioni del dispositivo e in particolare si concentra l'attenzione su:

- *Throughput totale* Th_{tot} = numero medio di pacchetti complessivamente trasmessi per slot.
- *Throughput locale* Th_w = numero medio di pacchetti originati dall'attività di sensing trasmessi per slot.

- *Throughput esogeno* Th_r = numero medio di pacchetti esogeni trasmessi per slot.

Naturalmente Th_w e Th_r compongono insieme Th_{tot} ovvero $Th_{tot} = Th_w + Th_r$.

Per tutti i risultati proposti in questa sezione si è assunto a titolo esemplificativo $E_{tx} = 4$. Conseguentemente, in base a quanto detto in sezione 5.1, si è supposto anche $E_{rx} = 4$ (dato che $K_{rx} = 1$), $E_{id} = 0$ (poiché $K_{id} = 0$) e $N = 4000$ (con $K_N = 1000$). Per quanto riguarda il livello di guardia della batteria si è posto $\rho = 0$ e quindi $N_{lim} = 0$. Infine, per la fonte di energia rinnovabile si è adottato il modello a due stati, valido specificatamente per l'energia solare, supponendo per la distribuzione $q_1(\cdot)$ un andamento a campana centrato sull'insieme di valori $\{1, 2, \dots, E_{max}\}$ con $\sigma = 1$.

Si sono quindi studiati gli effetti che le variazioni dei parametri rimanenti hanno sul throughput che il sensore radio è in grado di garantire seguendo la politica ottima. In ogni prova l'evoluzione del sistema è stata simulata per una durata di $2 \cdot 10^6$ slot.

5.2.1 Politica ottima e altre politiche

Il primo aspetto interessante da valutare è se la politica calcolata mediante l'ottimizzatore sia effettivamente la politica ottima voluta. Una volta forniti i parametri in ingresso, l'ottimizzatore calcola la politica ottima secondo il criterio del costo medio per slot per il processo decisionale rappresentativo del sensore radio. Se il modello proposto è ben definito, soprattutto per quanto riguarda i costi sul singolo slot, e la procedura di ottimizzazione è corretta, tale politica, una volta applicata al sistema, dovrebbe portare il dispositivo a massimizzare il throughput. Questo è quanto si vuole verificare in questa sezione: si valutano le prestazioni del sensore radio ottenute seguendo la politica ottima calcolata e le si confrontano con quelle relative ad altre politiche.

Nello specifico le politiche oggetto della comparazione sono tre:

- *Politica ottima* μ^* : fornita in uscita dall'ottimizzatore.
- *Politica random* μ_{rnd} : ottenuta scegliendo in modo completamente casuale l'azione da eseguire in ciascuno stato. Preso in considerazione un qualsiasi stato $i \in \mathcal{S}$, l'azione $\mu_{rnd}(i)$ viene selezionata secondo distribuzione uniforme nell'action set $\mathcal{A}(i)$, si assume, cioè, che $\mu_{rnd}(i)$ sia con la stessa probabilità uno degli elementi di $\mathcal{A}(i)$.

- *Politica greedy* μ_{grd} : qualsiasi sia lo stato occupato dal sistema all'inizio di un generico slot, viene scelta l'azione più conveniente *nel breve termine*, ovvero in quello stesso slot. In ogni stato viene eseguita quella che sarebbe l'azione migliore per massimizzare il throughput se la capacità della batteria fosse infinita e quindi non ci fossero vincoli sui consumi energetici. In sostanza, per ogni stato $i \in \mathcal{S}$ l'azione $\mu_{grd}(i)$ è quella che minimizza il costo $g(i, a)$ accumulato nel singolo slot corrente. Seguendo questa politica il sensore radio esegue in ogni slot l'azione che in quel dato momento porta ad accumulare il minor costo, ma si tratta di una politica "miope": non vengono considerati gli effetti che la scelta di tale azione ha sull'evoluzione futura del sistema e sui costi che verranno accumulati successivamente.

Per il confronto tra queste politiche si assume $K_{sx} = 1$, quindi $E_{sx} = 4$, e $\beta = 0.5$. Ciò significa imporre una condizione di assoluta equità tra l'attività di rilevazione del fenomeno di interesse e quella di inoltro di pacchetti esogeni: il valore fissato per il parametro β assicura che il sensore radio non attribuisca maggior rilevanza ad una delle due attività, mentre il valore assunto per K_{sx} fa sì che in ogni slot l'operazione di sensing e quella di ricezione portino al medesimo dispendio di energia (che è lo stesso anche per la trasmissione), cioè $E_{sx} = E_{rx} = E_{tx}$. Si suppone, poi, $K_{MAX} = 2$ e quindi $E_{MAX} = 8$: in un generico slot se la fonte di energia si trova nello stato $x = 0$ non fornisce alcuna unità di carica alla batteria; se, invece, lo stato occupato è $x = 1$ essa viene ricaricata con un numero di unità aleatorio secondo la distribuzione $q_1(\cdot)$ con andamento a campana centrato in $\{1, 2, \dots, 8\}$ con $\sigma = 1$.

Si simula l'evoluzione del sensore radio, in accordo con ciascuna delle tre politiche, al variare dell'intensità del traffico locale e di quello esogeno. In particolare si assume $P_{rx} = P_{sx}$ e si varia il valore di tali probabilità da 0 a 0.6, analizzando gli effetti a livello di throughput totale ottenuto.

Si valuta innanzitutto il caso in cui radio e sensore operino in modo indipendente. In figura 5.3 sono riportati i risultati relativi al throughput totale Th_{tot} , ottenuto adottando le tre diverse politiche, per i vari valori di $P_{rx} = P_{sx}$ considerati.

Come era lecito attendersi, qualsiasi sia la politica seguita dal dispositivo, il throu-

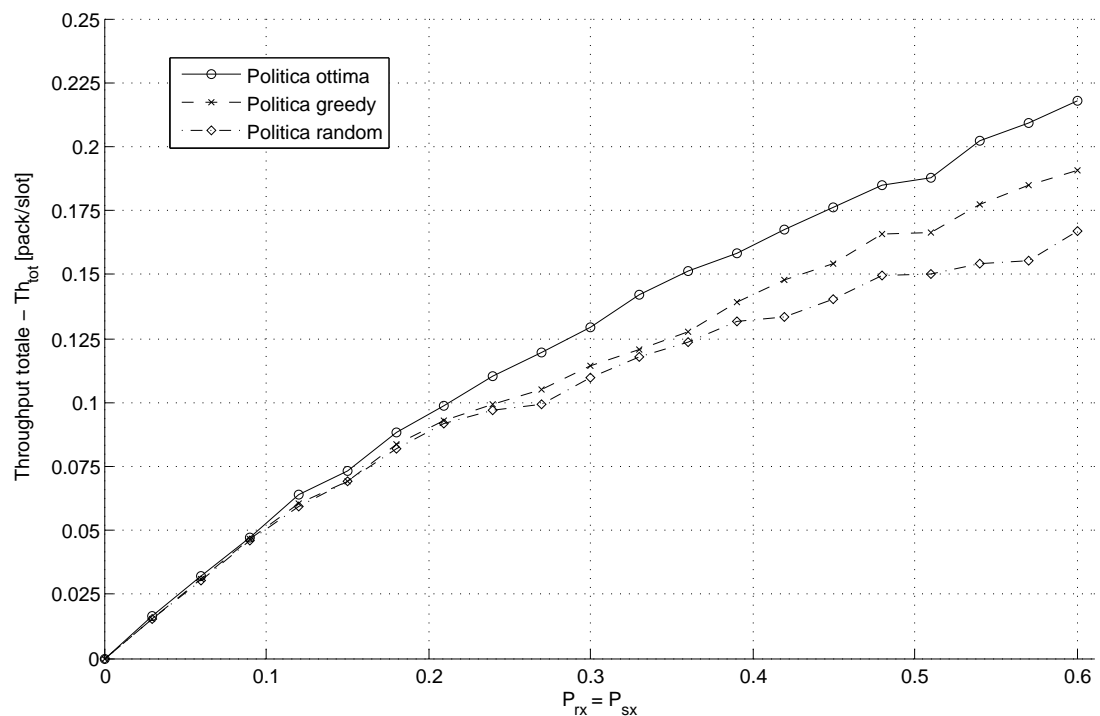


Figura 5.3 Throughput totale ottenuto adottando le tre diverse politiche al variare di P_{rx} e P_{sx} , con radio e sensore indipendenti.

ghput totale aumenta al crescere del traffico locale ed esogeno. All'aumentare di P_{rx} e P_{sx} , infatti, cresce la probabilità di acquisire con successo un nuovo pacchetto da trasmettere quando il dispositivo attiva la radio in ricezione o esegue l'attività di sensing mediante il sensore.

Per tutti i valori di $P_{rx} = P_{sx}$, poi, la politica ottima calcolata con l'ottimizzatore è effettivamente quella che consente di ottenere il throughput totale maggiore. Viceversa, come è ragionevole che sia, la politica random si rivela la meno conveniente delle tre, mentre la politica greedy garantisce prestazioni intermedie. Ciò testimonia che scegliere l'azione che appare migliore in un dato momento senza valutare quali siano gli effetti sull'evoluzione futura non porta a raggiungere prestazioni complessivamente ottime.

Del tutto analoghi sono i risultati che si ottengono nel caso si considerino radio e sensore mutuamente esclusivi, come evidenzia la figura 5.4. Anche in questo frangente, per tutti i valori di $P_{rx} = P_{sx}$ la politica che porta al throughput totale maggiore è quella calcolata dall'ottimizzatore.

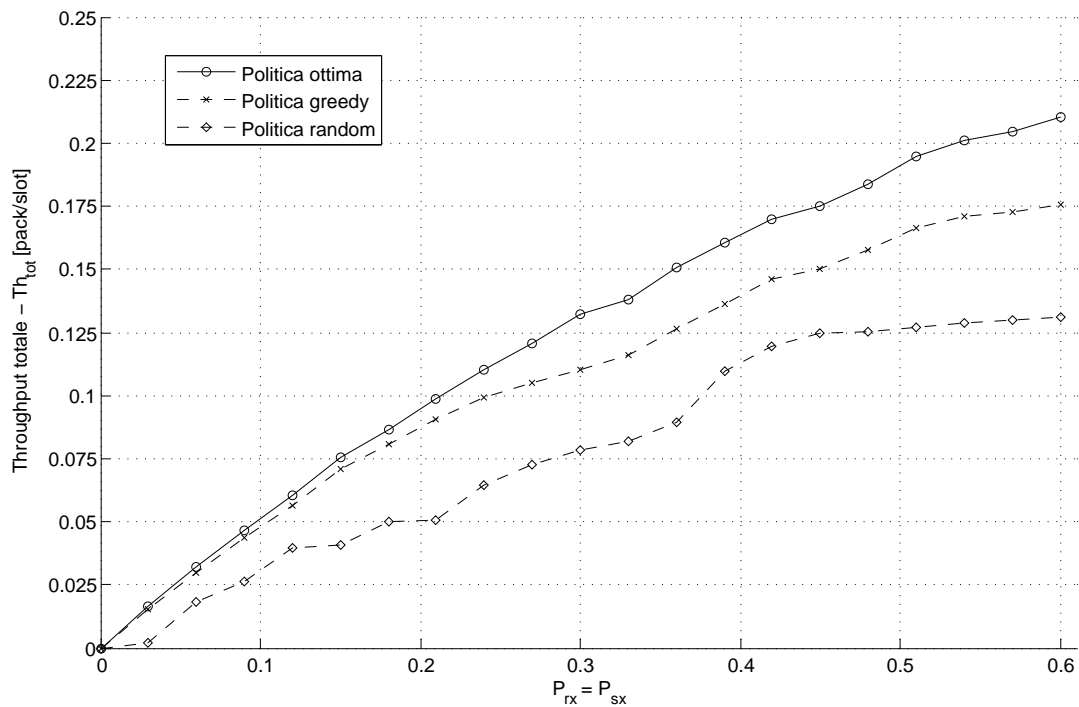


Figura 5.4 Throughput totale ottenuto adottando le tre diverse politiche al variare di P_{rx} e P_{sx} , con radio e sensore mutuamente esclusivi.

Nel complesso i risultati mostrati in figura 5.3 e in figura 5.4 dimostrano che il modello proposto nel capitolo 3 è ben definito e che l'algoritmo di ottimizzazione implementato nell'ottimizzatore opera correttamente, consentendo di determinare la politica di scheduling ottima dal punto di vista del throughput sia con radio e sensore indipendenti che mutuamente esclusivi.

5.2.2 Utilizzo del parametro β

Come discusso in sezione 3.6.4, il parametro $\beta \in [0, 1]$ viene introdotto nel modello del sensore radio per poter specificare quale attività tra rilevazione e inoltro sia più importante per il dispositivo considerato e in che misura. Se $\beta > 0.5$ ha maggior rilevanza l'attività di monitoraggio, dato che β rappresenta la probabilità che in un qualsiasi slot il sistema operi dando precedenza alla rilevazione del fenomeno di interesse; quanto più il valore di β è vicino a 1 tanto più marcata è la preferenza per questa attività. Per $\beta < 0.5$, invece, la situazione è speculare e l'attività prediletta è quella di inoltro.

In questa sezione l'interesse è quello di verificare che le variazioni del parametro β abbiano realmente l'effetto desiderato sulle prestazioni del dispositivo. Ci si aspetta che al variare del valore assunto da β i throughput Th_w e Th_r contribuiscano in percentuale diversa, proporzionale a β , al throughput totale Th_{tot} ottenuto simulando l'evoluzione del dispositivo secondo la politica ottima. In particolare, valori di β superiori a 0.5 dovrebbero favorire il throughput Th_w relativo ai pacchetti originati dall'attività di sensing, mentre per valori inferiori a 0.5 dovrebbe essere avvantaggiato Th_r .

I parametri del modello, tranne β , vengono fissati in modo da garantire una situazione di completa equità tra le due attività di rilevazione e inoltre: si pone $K_{sx} = 1$ in modo tale che $E_{sx} = E_{rx} = E_{tx}$ e $P_{rx} = P_{sx} = 0.3$ così che la probabilità di acquisire con successo un nuovo pacchetto risulti la stessa sia in ricezione che in sensing. Per la sorgente di energia, invece, si suppongono valide le stesse condizioni di sezione 5.2.1. Si assume, infine, che la radio e il sensore siano indipendenti, ma risultati analoghi si ottengono anche con radio e sensore mutuamente esclusivi. Con queste ipotesi, nel caso il sensore radio non prediliga nessuna delle due attività, è lecito attendersi che il contributo di Th_w e Th_r in Th_{tot} sia lo stesso. Questo risultato, dunque, si dovrebbe ottenere quando $\beta = 0.5$. Gli altri valori di β , invece, dovrebbero portare i throughput Th_w e Th_r ad essere tra loro diversi.

In figura 5.5 sono riportati i risultati relativi ai throughput Th_{tot} , Th_w , Th_r ottenuti simulando l'evoluzione del dispositivo in accordo con la politica ottima calcolata al variare di β da 0 a 1. In particolare, in figura 5.5(a) viene mostrato l'andamento dei tre throughput in funzione di β ; in figura 5.5(b), invece, viene evidenziato, per ciascun valore di β considerato, il contributo percentuale di Th_w e Th_r in Th_{tot} .

Osservando la figura 5.5(a), si può notare come il throughput totale Th_{tot} rimanga sostanzialmente costante indipendentemente dal valore di β . Ciò che cambia è la composizione di Th_{tot} , cioè il valore delle sue componenti Th_w e Th_r : come ci si attendeva, al crescere di β da 0 a 1, il throughput Th_r relativo ai pacchetti esogeni diminuisce progressivamente lasciando spazio a Th_w che diventa, invece, predominante.

L'effetto della variazione del parametro β sui throughput Th_w e Th_r è ancora più evidente nella figura 5.5(b). Quando $\beta = 0$ è solo Th_r a contribuire al throughput totale,

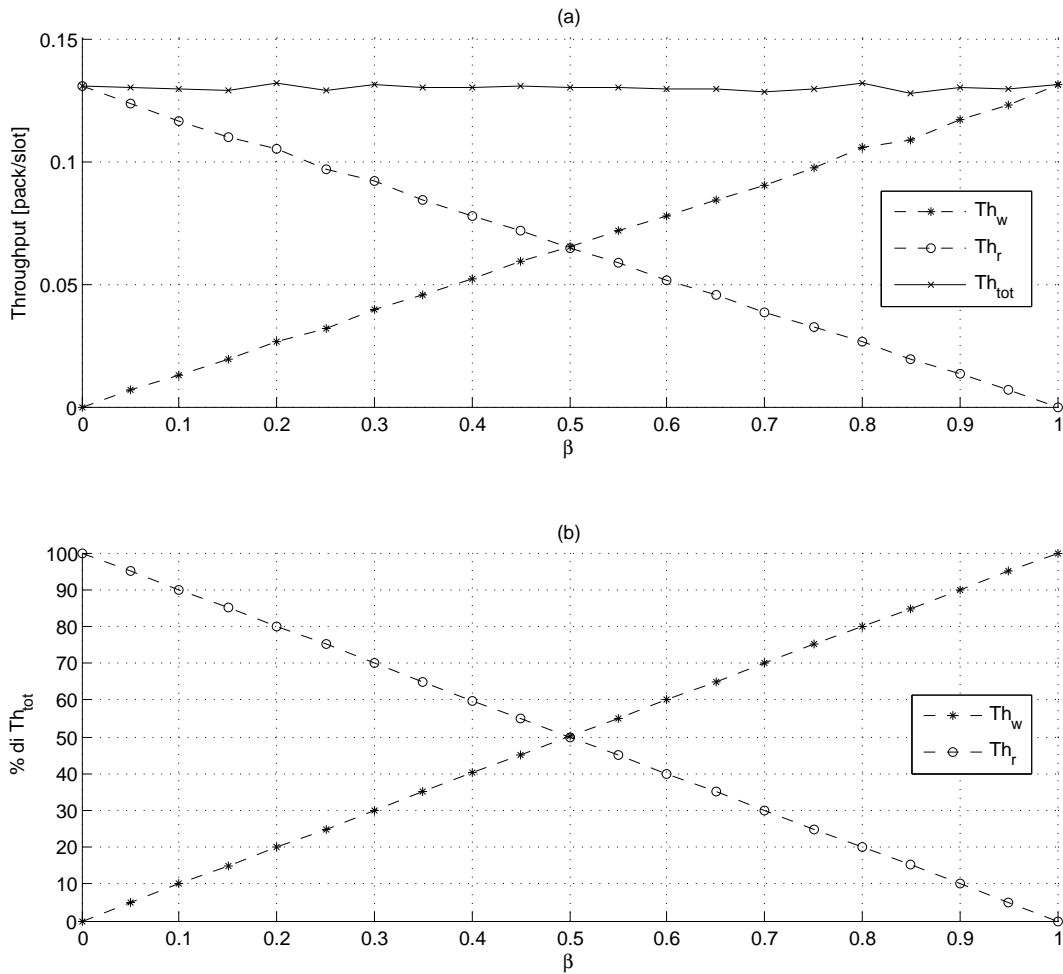


Figura 5.5 (a) Andamento di Th_{tot} , Th_w e Th_r al variare di β . (b) Contributo percentuale di Th_w e Th_r in Th_{tot} al variare di β .

come è ragionevole che sia dato che il dispositivo opera in tutti gli slot preferendo l'attività di inoltro. La situazione è invece speculare per $\beta = 1$ dove il throughput totale è interamente costituito da Th_w . Per $\beta = 0.5$, poi, il risultato è proprio quello voluto: i throughput Th_w e Th_r contribuiscono in egual misura al throughput totale. Infine, gli altri valori di β portano a risultati intermedi in cui Th_{tot} è composto da Th_w e Th_r secondo percentuali proporzionali a β .

Questi risultati confermano che, agendo sul parametro β , le conseguenze sulle prestazioni del sensore radio sono proprio quelle desiderate. Tale parametro può essere efficacemente utilizzato come “manopola” per adattare le prestazioni del dispositivo a seconda del ruolo che esso ricopre nella rete in cui si trova.

5.2.3 Consumo energetico dell'attività di sensing

Nei risultati finora considerati si è supposto che il dispendio di energia necessario per compiere la ricezione o il sensing di un pacchetto fosse il medesimo e pari a quello richiesto per la trasmissione, cioè $E_{sx} = E_{rx} = E_{tx}$. In realtà, come già detto in sezione 5.1.1, il consumo energetico dovuto all'attività di sensing può essere inferiore o anche superiore a quello della radio in ricezione o trasmissione. Ciò dipende dal particolare dispositivo utilizzato e dalla specifica applicazione. Per questo motivo, in questa sezione si valuta come varino le prestazioni del sensore radio per valori di $K_{sx} \neq 1$, quindi per $E_{sx} \neq E_{rx}$. Si ricorda, in particolare, che quando $K_{sx} < 1$ allora $E_{sx} < E_{tx}$, mentre se $K_{sx} > 1$ si ha $E_{sx} > E_{tx}$.

A titolo esemplificativo si considerano ancora radio e sensore indipendenti, essendo analoga la situazione per radio e sensore mutuamente esclusivi. Per la fonte di energia rinnovabile si mantengono le condizioni delle sezioni precedenti, ponendo in particolare $K_{MAX} = 2$ (quindi $E_{MAX} = 8$). Inoltre si fissano $P_{rx} = P_{sx} = 0.3$ e $\beta = 0.5$. Si suppone, quindi, che K_{sx} possa variare da 0.25 a 1.5 con passo 0.25, ovvero si considerano per E_{sx} i valori $\{1, 2, \dots, 6\}$. Per ciascun valore di K_{sx} , e quindi di E_{sx} , si calcola la politica ottima e in accordo con quest'ultima si simula l'evoluzione del sensore radio analizzando i throughput ottenuti.

La figura 5.6 riassume i risultati delle simulazioni eseguite. In particolare, in figura 5.6(a) è mostrato l'andamento dei throughput Th_{tot} , Th_w e Th_r al variare di K_{sx} , mentre la figura 5.6(b) mette in evidenza il contributo percentuale di Th_w e Th_r in Th_{tot} in corrispondenza dei diversi valori di K_{sx} considerati.

Innanzitutto si noti, in figura 5.6(a), l'andamento del throughput totale Th_{tot} in funzione di K_{sx} : all'aumentare di K_{sx} vi è una diminuzione progressiva di Th_{tot} . Questo è dovuto al fatto che al crescere di K_{sx} l'energia necessaria per compiere una rilevazione aumenta e conseguentemente, a parità di azioni eseguite dal dispositivo, l'energia consumata è progressivamente più grande. I vincoli dovuti alla capacità finita della batteria si fanno dunque più stringenti, limitando maggiormente le prestazioni del sistema.

Ancora dalla figura 5.6(a) si evince che al crescere di K_{sx} la flessione più evidente

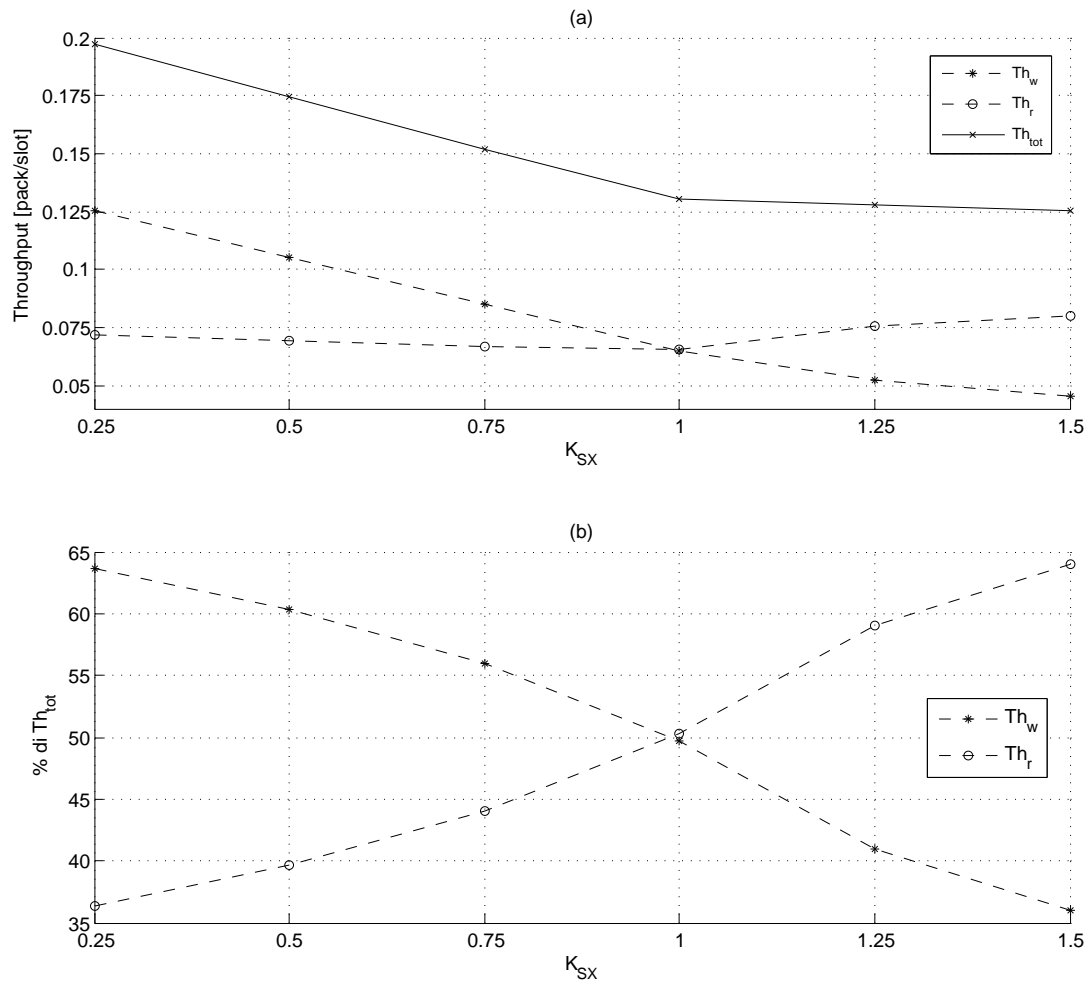


Figura 5.6 (a) Andamento di Th_{tot} , Th_w e Th_r al variare di K_{sx} . (b) Contributo percentuale di Th_w e Th_r in Th_{tot} al variare di K_{sx} .

riguarda Th_w : un dispendio di energia più consistente per compiere l'operazione di sensing costringe il sensore radio ad eseguire meno azioni di questo tipo e quindi a trasmettere mediamente meno pacchetti contenenti i dati sul fenomeno misurato. Al variare di K_{sx} da 0.25 a 1 si nota anche una leggero calo per quanto riguarda Th_r : avere a disposizione complessivamente meno energia a causa dell'attività di sensing più dispendiosa, porta indirettamente il dispositivo a ridurre anche l'attività di ricezione e conseguente trasmissione dei pacchetti esogeni. Per $K_{sx} > 1$, poi, si ha $Th_r > Th_w$ in quanto l'attività di inoltro diventa più conveniente, dal punto di vista energetico, rispetto a quella di rilevazione.

In figura 5.6(b), poi, viene mostrato come il throughput totale Th_{tot} si distribuisca tra Th_w e Th_r al variare di K_{sx} . Come era ragionevole attendersi, per valori di $K_{sx} < 1$ è Th_w ad essere predominante: dato che $P_{rx} = P_{sx}$ e $\beta = 0.5$, il minor consumo energetico dell'operazione di sensing rispetto a quella di ricezione rende più attraente l'attività di monitoraggio piuttosto che quella di inoltro. Quando $K_{sx} > 1$ la situazione è naturalmente invertita.

5.2.4 Energia fornita dalla fonte rinnovabile

Come si è già osservato, il numero massimo di unità di carica E_{MAX} che la fonte di energia rinnovabile è in grado di fornire alla batteria in uno slot dipende dalle particolari condizioni operative e dal dispositivo considerato. In questa sezione, dunque, si analizzano le prestazioni del sensore radio in funzione di E_{MAX} . A tal fine, si agisce sul parametro K_{MAX} ricordando che $E_{MAX} = K_{MAX}E_{tx}$.

Si impongono delle condizioni di assoluta equità tra l'attività di rilevazione e quella di inoltro fissando $\beta = 0.5$, $K_{sx} = 1$ e $P_{rx} = P_{sx} = 0.3$. Si suppone, poi, che K_{MAX} possa assumere valori da 0.25 a 2 con passo 0.25, ovvero si assume che i valori ammissibili per E_{MAX} siano $\{1, 2, \dots, 8\}$. Come per tutti i risultati visti finora, si rappresenta la sorgente di energia mediante il modello a due stati, assumendo per la distribuzione $q_1(\cdot)$ un andamento a campana centrato sull'insieme di valori $\{1, 2, \dots, E_{MAX}\}$ con $\sigma = 1$. Anche in questo caso, senza perdita di generalità, si considerano radio e sensore indipendenti dato che dal punto di vista qualitativo i risultati sono analoghi nel caso in cui radio e sensore siano mutuamente esclusivi.

In figura 5.7 sono riportati i valori di Th_{tot} , Th_w e Th_r ottenuti simulando l'evoluzione del sensore radio in accordo con la politica ottima calcolata per i diversi valori di K_{MAX} . Il risultato è piuttosto evidente: all'aumentare di K_{MAX} il throughput totale Th_{tot} cresce progressivamente e, naturalmente, anche le sue componenti Th_w e Th_r seguono lo stesso andamento. Per valori crescenti di K_{MAX} , infatti, aumenta l'apporto energetico su cui il sensore radio può contare in ciascuno slot. Conseguentemente il dispositivo può essere più aggressivo nell'utilizzo dell'energia contenuta nella batteria e raggiungere, così, prestazioni migliori. I throughput Th_w e Th_r presentano un an-

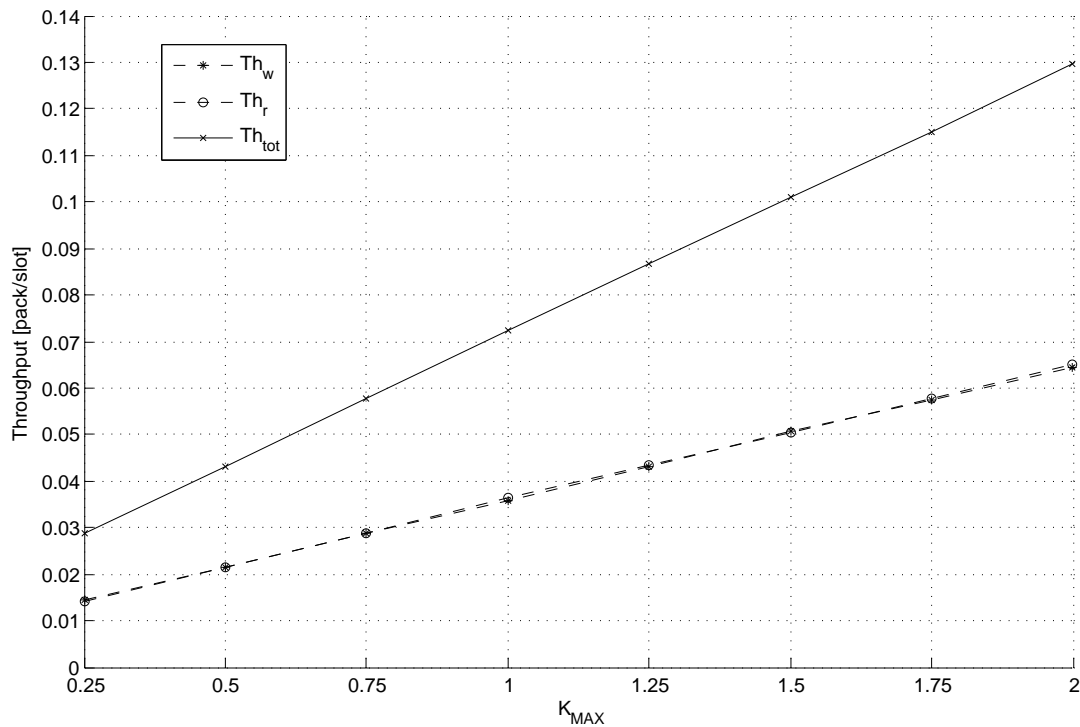


Figura 5.7 Andamento di Th_{tot} , Th_w e Th_r al variare di K_{MAX} .

damento praticamente identico poiché le condizioni imposte assicurano una completa equità tra le attività di rilevazione e inoltro.

5.3 Risultati ottenuti mediante analisi delle politiche ottime

I risultati presentati finora sono stati ottenuti simulando l'evoluzione del sensore radio, in accordo con la politica ottima restituita dall'ottimizzatore in base ai parametri stabiliti, e valutando, poi, le prestazioni raggiunte. In questa sezione, invece, l'intenzione è quella di studiare direttamente alcune caratteristiche della politica ottima calcolata e valutare come queste siano influenzate dalla scelta dei parametri, ovvero dalle condizioni operative in cui si trova il dispositivo considerato.

Il procedimento seguito per l'analisi delle politiche ottime è il seguente: innanzitutto si fissano i parametri del modello e si forniscono in ingresso all'ottimizzatore; quest'ultimo viene poi utilizzato per determinare la politica ottima nelle condizioni poste; infine si valuta quanto è attivo il sensore radio, secondo la politica ottima, al

variare della quantità residua di energia disponibile, ovvero si studia il livello di attività del dispositivo in funzione dello stato di carica della batteria.

A tal fine, l'insieme dei valori ammissibili $\{0, 1, 2, \dots, N\}$ per la variabile n , che indica il numero di unità di carica rimaste nella batteria, viene partizionato in un numero l di sottoinsiemi, ciascuno dei quali è rappresentativo di un diverso livello di carica. Per chiarezza, si assuma $N = 12000$ e $l = 20$: l'insieme $\{0, 1, 2, \dots, 12000\}$ viene suddiviso nei sottoinsiemi $\Delta N_0 = \{0, 1, \dots, 599\}$, $\Delta N_1 = \{600, 601, \dots, 1199\}$ fino a giungere a ΔN_{l-1} . In questo modo, il sottoinsieme ΔN_0 , ad esempio, rappresenta un livello di carica della batteria compreso tra lo 0% e il 5% della capacità totale, il sottoinsieme ΔN_1 si riferisce all'intervallo 5% – 10% e così via.

Per ciascun sottoinsieme ΔN_j si considerano gli stati $i = (n, x, w, r, v)$ tali che $n \in \Delta N_j$. Questi sono tutti gli stati del sistema in cui il livello di carica della batteria è quello rappresentato da ΔN_j . Si valuta, allora, in quanti di questi stati la politica ottima indica che il sensore radio è attivo. Nello specifico, si calcola la percentuale di stati in cui l'azione ottima prevede che il dispositivo non rimanga a riposo, cioè in cui l'azione ottima è diversa da (OFF, ID) nel caso di radio e sensore indipendenti e da (ID) con radio e sensore mutuamente esclusivi. In questo modo si ottiene un'importante indicazione su quanto deve essere attivo il sensore radio secondo la politica ottima in funzione dello stato di carica della batteria. Se per un certo livello di carica la percentuale calcolata è alta significa che il dispositivo può essere aggressivo nell'utilizzo della batteria, viceversa, se la percentuale è bassa, il sensore radio deve essere più parsimonioso.

Per giungere ai risultati proposti in questa sezione viene condotta l'analisi appena descritta sulle politiche ottime ottenute per diversi valori dei parametri del modello. Particolarmente interessante è l'introduzione di alcune variazioni sulla descrizione statistica della fonte di energia rinnovabile.

5.3.1 Modello a stato singolo per la fonte di energia rinnovabile

Si assuma di adottare per la fonte di energia rinnovabile il modello a stato singolo con $X = 0$. In questo caso, come già detto più volte, la sorgente è completamente rappresen-

tata dalla distribuzione $q_0(\cdot)$ definita sull'insieme di valori $\Delta E_0 = \{0, 1, \dots, E_{MAX}\}$. Il numero di unità di carica acquisito dalla batteria in ciascuno slot segue tale distribuzione. L'interesse è quello di studiare come varia il comportamento ottimo per il sensore radio, in corrispondenza dei diversi livelli di carica della batteria, a seconda del tipo di distribuzione $q_0(\cdot)$ considerato.

Per quanto riguarda i vari parametri del modello si suppone, a titolo esemplificativo, di fissare $E_{tx} = 12$ e conseguentemente $N = 12000$ (dato che $K_N = 1000$), $E_{rx} = 12$ e $E_{id} = 0$ (con $K_{rx} = 1$ e $K_{id} = 0$). Si impongono, poi, condizioni di completa equità tra l'attività di rilevazione e quella di inoltro ponendo $K_{sx} = 1$, cioè $E_{sx} = 12$, $\beta = 0.5$ e $P_{rx} = P_{sx} = 0.15$. Per il limite di guardia della batteria si considera $\rho = 0$ e quindi $N_{lim} = 0$. Infine, per la fonte di energia rinnovabile si fissa $K_{MAX} = 2/3$, cioè $E_{MAX} = 8$, e si assume per la distribuzione $q_0(\cdot)$ un andamento a campana centrato su $\{0, 1, \dots, 8\}$ con σ variabile. In particolare, si fa riferimento alle diverse distribuzioni mostrate in figura 5.2.

Per ciascun valore di σ si calcola la relativa politica ottima e si valuta quale sia, secondo tale politica, il livello di attività del sensore radio per i diversi stati di carica della batteria. I risultati ottenuti, nel caso di radio e sensore indipendenti, sono illustrati in figura 5.8.

Innanzitutto si osserva che, indipendentemente dalla particolare distribuzione $q_0(\cdot)$ considerata, le curve presentano un andamento analogo, dal punto di vista qualitativo, e abbastanza prevedibile: quando la batteria è molto carica la percentuale di stati in cui la politica ottima prevede che il sensore radio sia attivo è elevata; al decrescere della carica residua, invece, la percentuale diminuisce progressivamente. Nelle particolari condizioni operative poste, si nota che quando la batteria è carica per almeno il 60% della sua capacità, in qualsiasi stato si trovi il sistema e per tutte le distribuzioni $q_0(\cdot)$ considerate, la politica ottima prevede che il dispositivo esegua una qualche attività. Un'indicazione di questo tipo è molto utile: finché il livello di carica della batteria è al di sopra di tale limite, il sensore radio può permettersi un comportamento aggressivo, eseguendo, quando possibile, tutte le attività che contribuiscono a massimizzare le prestazioni.

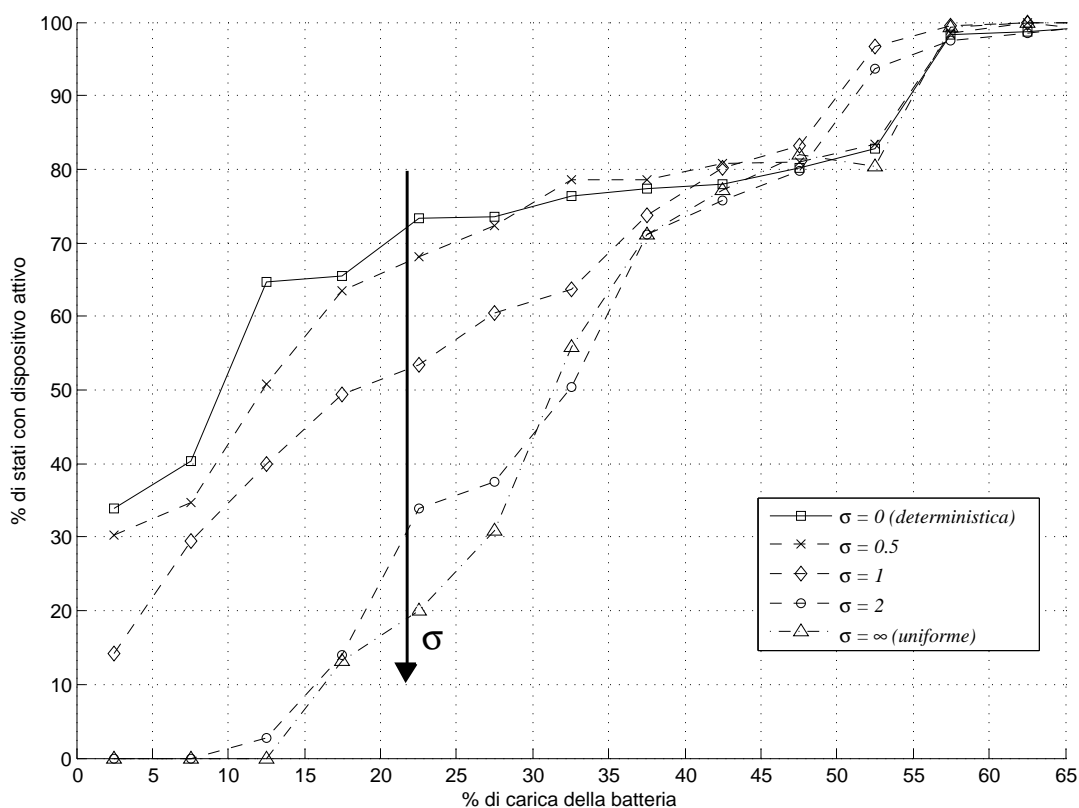


Figura 5.8 Livello di attività del sensore radio, secondo la politica ottima, in funzione dello stato di carica della batteria e al variare di $q_0(\cdot)$.

La particolare distribuzione $q_0(\cdot)$ considerata influisce in modo significativo sulla politica ottima per bassi livelli di carica della batteria, ovvero quando l'attività del sensore radio è più vincolata dal punto di vista energetico.

Nello specifico caso in esame, si notano importanti differenze tra le politiche ottime, ottenute variando $q_0(\cdot)$, quando il livello di carica della batteria è inferiore al 40% della capacità totale. In tali condizioni, come si vede in figura 5.8, al crescere di σ il comportamento ottimo per il sensore radio è sempre più prudente: per $\sigma = 0$, cioè con $q_0(\cdot)$ deterministica, la politica ottima porta il dispositivo ad essere più attivo rispetto a quanto avviene per valori superiori di σ ; d'altra parte con $\sigma = \infty$, ovvero per $q_0(\cdot)$ uniforme, la politica ottima induce, invece, un utilizzo molto più parsimonioso dell'energia rispetto agli altri casi. Quando $\sigma = 0$ e la distribuzione $q_0(\cdot)$ è deterministica, il sensore radio può contare in ciascuno slot su un apporto energetico

certo, ovvero il numero di unità di carica che vengono fornite alla batteria è noto senza alcuna incertezza. Al crescere di σ , fino al caso limite $\sigma = \infty$, l'incertezza sul numero di unità che in ogni slot ricaricano la batteria aumenta progressivamente, diventando massima quando la distribuzione $q_0(\cdot)$ è uniforme. La politica ottima viene calcolata tenendo conto di questo grado di incertezza e conseguentemente ne è influenzata. Se in ogni slot il contributo energetico della fonte è certo, il sensore radio può assumersi il rischio di eseguire alcune attività anche quando la batteria è molto scarica, dato che conosce con sicurezza come e quanto sarà ricaricata negli slot successivi. Se, invece, l'incertezza è grande, il dispositivo deve utilizzare l'energia in modo molto più accorto poiché non è noto in quanto tempo verrà recuperata l'energia eventualmente consumata.

5.3.2 Livello di guardia della batteria

Fino a questo momento si è sempre assunto come limite di guardia della batteria $N_{lim} = 0$, consentendo al sensore radio di utilizzare senza problemi tutta l'energia a disposizione. Si vogliono valutare, a questo punto, gli effetti che si hanno sulla politica ottima variando il limite di guardia.

Si pongono le stesse condizioni operative viste in sezione 5.3.1. In particolare, per la fonte di energia rinnovabile si adotta il modello a stato singolo con distribuzione $q_0(\cdot)$ uniforme ($\sigma = \infty$) sull'insieme di valori ammissibili $\{0, 1, \dots, 8\}$. Per quanto riguarda il parametro ρ , si considerano i valori $\{0, 0.2, 0.3, 0.4\}$, dunque N_{lim} si assume variabile nell'insieme $\{0, 2400, 3600, 4800\}$ (ricordando che $N = 12000$). La radio e il sensore, poi, vengono supposti indipendenti.

In figura 5.9 vengono proposti i risultati ottenuti relativamente al livello di attività del dispositivo, in funzione dello stato di carica della batteria, secondo le politiche ottime calcolate per i diversi valori di ρ .

Come è possibile notare, variando il livello di guardia della batteria si ottiene l'effetto desiderato: la politica ottima porta il dispositivo ad essere attivo solo quando lo stato di carica della batteria è sopra la soglia fissata, quindi, al crescere di ρ , il sistema inizia ad essere attivo a partire da livelli di carica sempre maggiori. In tutti i casi considerati, quando il numero di unità di carica residue è inferiore a N_{lim} , la politica ottima indica al

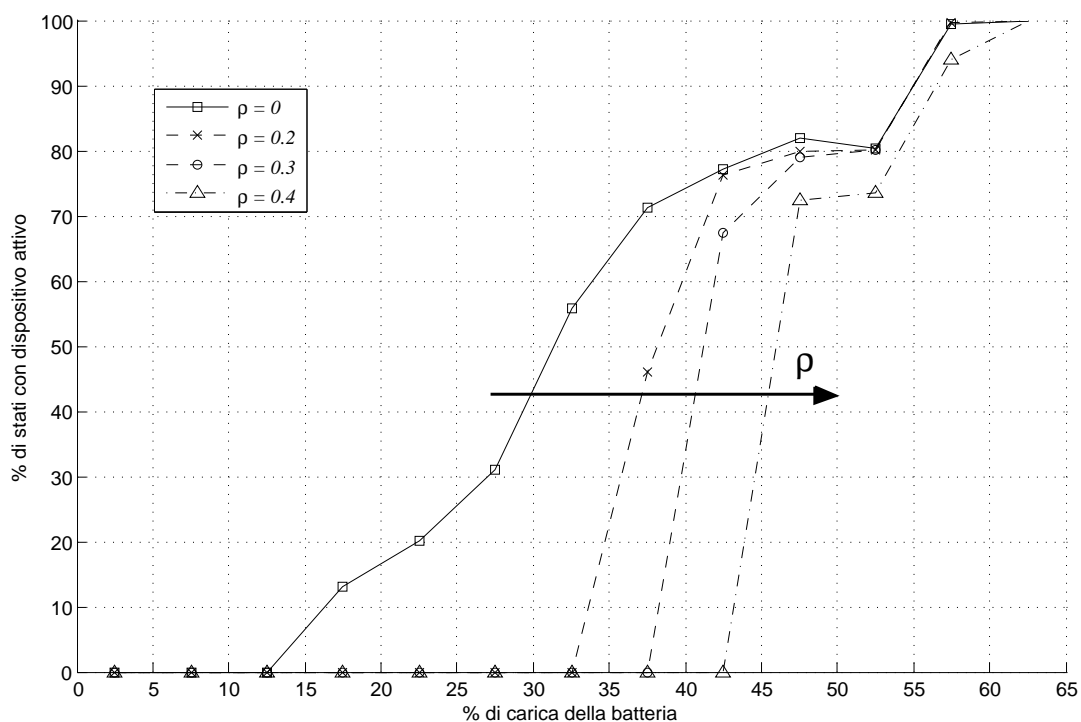


Figura 5.9 Livello di attività del sensore radio, secondo la politica ottima, in funzione dello stato di carica della batteria e al variare di ρ .

sensore radio di rimanere a riposo in attesa che la batteria ritorni sopra il limite imposto. Pertanto, come si voleva, la politica ottima impedisce che il dispositivo scarichi la batteria oltre il livello desiderato e fissato mediante il parametro ρ .

5.3.3 Modalità di funzionamento

Anziché imporre condizioni di assoluta equità tra l'attività di rilevazione e quella di inoltro, si suppone ora che il consumo energetico per eseguire l'operazione di sensing sia inferiore rispetto a quello dovuto alla ricezione e alla trasmissione. Si studiano, quindi, le conseguenze che tale assunzione ha sulla politica ottima.

Nello specifico, ai parametri del modello vengono assegnati gli stessi valori assunti in sezione 5.3.1. Ciò che cambia sono i parametri K_{sx} , che viene fissato a 0.5, e ρ , che viene posto pari a 0.2. Conseguentemente si assume che il numero di unità di carica E_{sx} richiesto per il sensing di un pacchetto sia la metà di $E_{rx} = 12$ e che il livello di guardia della batteria sia $N_{lim} = 2400$. In più, la distribuzione $q_0(\cdot)$ viene supposta

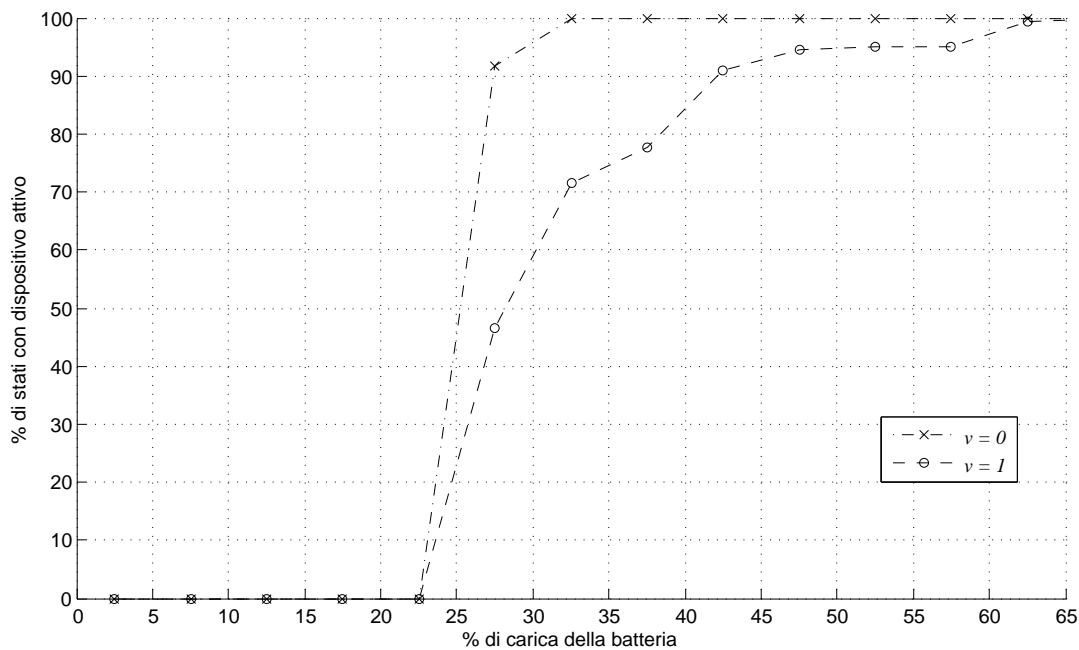


Figura 5.10 Livello di attività del sensore radio, secondo la politica ottima, in funzione dello stato di carica della batteria, per $v = 0$ e $v = 1$.

uniforme sull'insieme di valori ammissibili $\{0, 1, \dots, 8\}$ ($\sigma = \infty$). Anche in questo caso, poi, radio e sensore si considerano indipendenti.

L'aspetto che si intende considerare è sempre il livello di attività del dispositivo, secondo la politica ottima, in funzione dello stato della batteria. Tuttavia, questa volta, per ciascun livello di carica è interessante valutare separatamente gli stati in cui $v = 0$ e quelli in cui $v = 1$. Nel caso in cui $v = 0$, infatti, il sensore radio opera privilegiando l'attività di rilevazione, la quale può trarre beneficio dal minor dispendio energetico necessario per la misurazione del fenomeno di interesse. Quando $v = 1$, invece, il dispositivo predilige l'attività di inoltro che, complessivamente, è più onerosa dal punto di vista dei consumi. I risultati ottenuti dall'analisi della politica ottima, calcolata in queste condizioni operative, sono proposti in figura 5.10.

Si può notare come la politica ottima porti il dispositivo ad essere in generale più attivo negli stati in cui $v = 0$ rispetto a quelli in cui $v = 1$. La differenza diventa tanto più evidente quanto più ci si avvicina al livello di guardia imposto per la batteria. Ciò è dovuto al fatto che l'attività di inoltro è nel complesso più dispendiosa in termini

energetici. Pertanto quando il numero di unità di carica a disposizione si avvicina al limite N_{lim} , la politica ottima induce il sensore radio ad essere maggiormente attivo quando si dedica all'attività di rilevazione piuttosto che a quella di inoltro. Nel primo caso, infatti, l'energia spesa è complessivamente minore e può dunque essere recuperata più velocemente che nel secondo caso. Ecco spiegato il diverso livello di attività del dispositivo nei casi $v = 0$ e $v = 1$ soprattutto in prossimità della soglia di guardia.

5.3.4 Modello a due stati per la fonte di energia rinnovabile

Si considera, a questo punto, una fonte di energia rinnovabile che alterna fasi in cui non fornisce alcuna unità di carica alla batteria con altre in cui la alimenta con un contributo energetico variabile. In particolare, si fa riferimento all'energia solare il cui andamento è proprio di questo tipo: nelle ore notturne l'apporto di energia è nullo, mentre durante il dì è più o meno significativo a seconda dell'intensità della radiazione. Per rappresentare tale sorgente di energia il modello adottato è quello a due stati ($X = 1$) con le assunzioni poste in sezione 5.1.3. Lo stato $x = 0$ rappresenta le ore notturne mentre lo stato $x = 1$ quelle diurne.

Quando la fonte di energia presenta queste caratteristiche, è interessante studiare quale sia secondo la politica ottima il livello di attività del dispositivo, in funzione dello stato di carica della batteria, distinguendo gli stati in cui $x = 0$ da quelli con $x = 1$. Quando $x = 1$, infatti, il sensore radio sa di poter contare su un contributo energetico non nullo da parte della sorgente. Viceversa, se $x = 0$ la fonte di energia è in una fase in cui non fornisce alcuna unità di carica alla batteria e dunque il dispositivo è consapevole di dover agire in modo autonomo, senza che la batteria possa essere ricaricata.

Quest'analisi viene condotta supponendo $E_{tx} = 8$, quindi $N = 8000$, $E_{rx} = 8$ e $E_{id} = 0$. Si impongono, poi, condizioni di equità tra l'attività di rilevazione ed inoltro assumendo $K_{sx} = 1$ (dunque $E_{sx} = 8$), $\beta = 0.5$ e $P_{rx} = P_{sx} = 0.3$. Il livello di guardia della batteria viene fissato ponendo $\rho = 0$, cioè $N_{lim} = 0$. Infine, per quanto riguarda la fonte di energia rinnovabile, si considera $K_{MAX} = 1$, quindi $E_{MAX} = 8$, e si adotta, come detto, il modello a due stati con $q_0(0) = 1$ e $q_1(\cdot)$ uniforme sull'insieme di valori ammissibili $\Delta E_1 = \{1, 2, \dots, 8\}$. I risultati ottenuti analizzando la politica

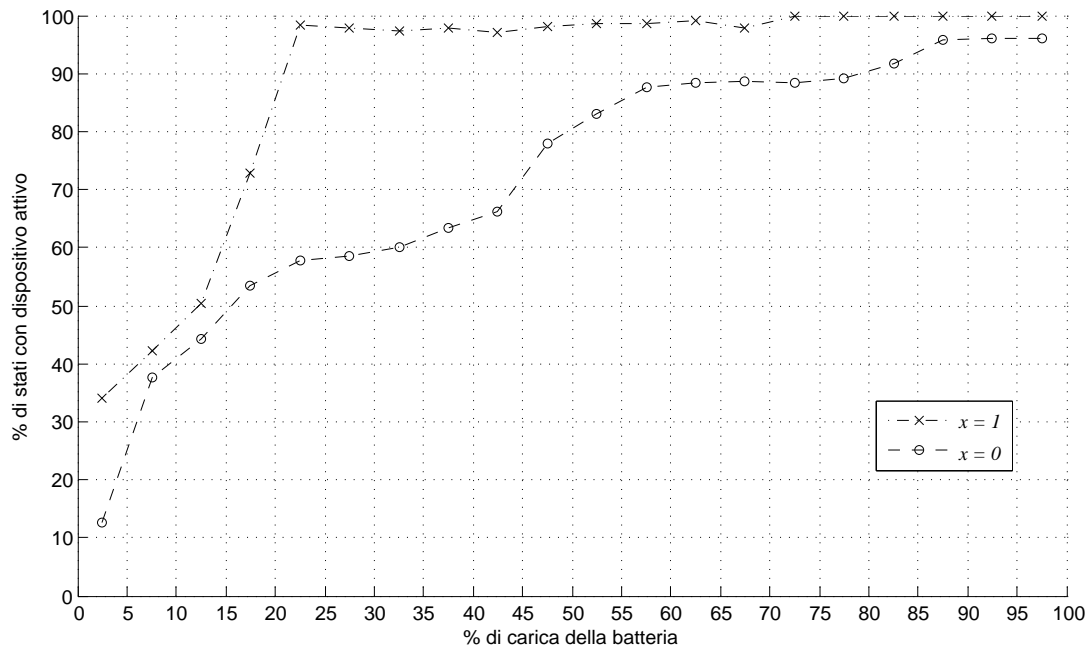


Figura 5.11 Livello di attività del sensore radio, secondo la politica ottima, in funzione dello stato di carica della batteria, per $x = 0$ e $x = 1$.

ottima, calcolata con queste ipotesi operative e con radio e sensore indipendenti, sono proposti in figura 5.11.

L'esito della prova è coerente con quanto ci si aspettava di ottenere: la politica ottima spinge il sensore radio ad essere più attivo negli stati in cui $x = 1$ rispetto a quelli in cui $x = 0$. Infatti, è ragionevole che il comportamento del dispositivo possa essere più aggressivo nei confronti della batteria durante le ore diurne, quando l'energia consumata può essere eventualmente recuperata grazie all'apporto della fonte rinnovabile. Durante la notte, invece, è lecito che il sensore radio debba comportarsi in modo più prudente dato che la batteria non può essere ricaricata. Naturalmente, la differenza tra le due curve in figura 5.11 è più evidente quando il numero di unità di carica rimaste è basso, cioè quando la situazione della batteria è più critica.

5.3.5 Modello a due stati vs modello a stato singolo

Quando la fonte di energia rinnovabile ha un comportamento a due fasi, come nel caso dell'energia solare, è utile rappresentarla mediante il modello a due stati: uno stato,

$x = 0$, per la fase in cui il contributo energetico della sorgente è nullo o scarso, e un altro, $x = 1$, per la fase durante la quale la fonte è in grado di ricaricare in modo più o meno intenso la batteria. Utilizzare tale modello, infatti, consente di calcolare la politica ottima per il dispositivo *tenendo conto* dello specifico comportamento della sorgente, cioè dell'alternanza tra le due fasi e della correlazione che vi è tra esse. Come verrà messo in evidenza in questa sezione, la politica ottima così calcolata è diversa da quella che si otterrebbe considerando separatamente le due fasi e adottando per ciascuna di esse il modello a stato singolo.

Per chiarezza, nel seguito si fa esplicitamente riferimento all'energia solare come fonte di energia considerata. Per i parametri del modello, esclusi quelli relativi alla fonte rinnovabile, si assumono validi gli stessi valori fissati in sezione 5.3.4, mentre la radio e il sensore questa volta si considerano mutuamente esclusivi.

Posto $K_{MAX} = 1$ e quindi $E_{MAX} = 8$, si assume che il numero di unità di carica per slot che la sorgente di energia è in grado di fornire alla batteria durante le ore diurne segua la distribuzione $q'(\cdot)$, uniforme sull'insieme di valori $\{2, 3, \dots, 8\}$. L'apporto energetico notturno, invece, si suppone in accordo con la distribuzione $q''(\cdot)$ uniforme su $\{0, 1\}$.

Adottando il modello a due stati, con lo stato $x = 0$ per le ore notturne e lo stato $x = 1$ per quelle diurne, una sorgente di energia di questo tipo può essere facilmente rappresentata ponendo $\Delta E_0 = \{0, 1\}$, $\Delta E_1 = \{2, 3, \dots, 8\}$ e definendo $q_1(\cdot) = q'(\cdot)$ e $q_0(\cdot) = q''(\cdot)$. Assumendo di utilizzare questo modello, mostrato in figura 5.12(a), si indica con μ^* la corrispondente politica ottima calcolata mediante l'ottimizzatore.

Si può pensare, però, di procedere anche in un altro modo: si possono considerare separatamente la fase diurna e quella notturna, utilizzando due modelli distinti per rappresentare la fonte rinnovabile durante il dì e nel corso della notte, e calcolare due diverse politiche ottime per le due fasi. Nello specifico, per descrivere il comportamento della sorgente durante le ore notturne è possibile adottare il modello a stato singolo ($X = 0$) proposto in figura 5.12(b), con $\Delta E_0 = \{0, 1\}$ e $q_0(\cdot) = q''(\cdot)$ uniforme sull'insieme di valori $\{0, 1\}$. Supponendo di utilizzare questo modello, si può allora calcolare la relativa politica ottima μ^0 adatta per la fase notturna. Allo stesso modo, per

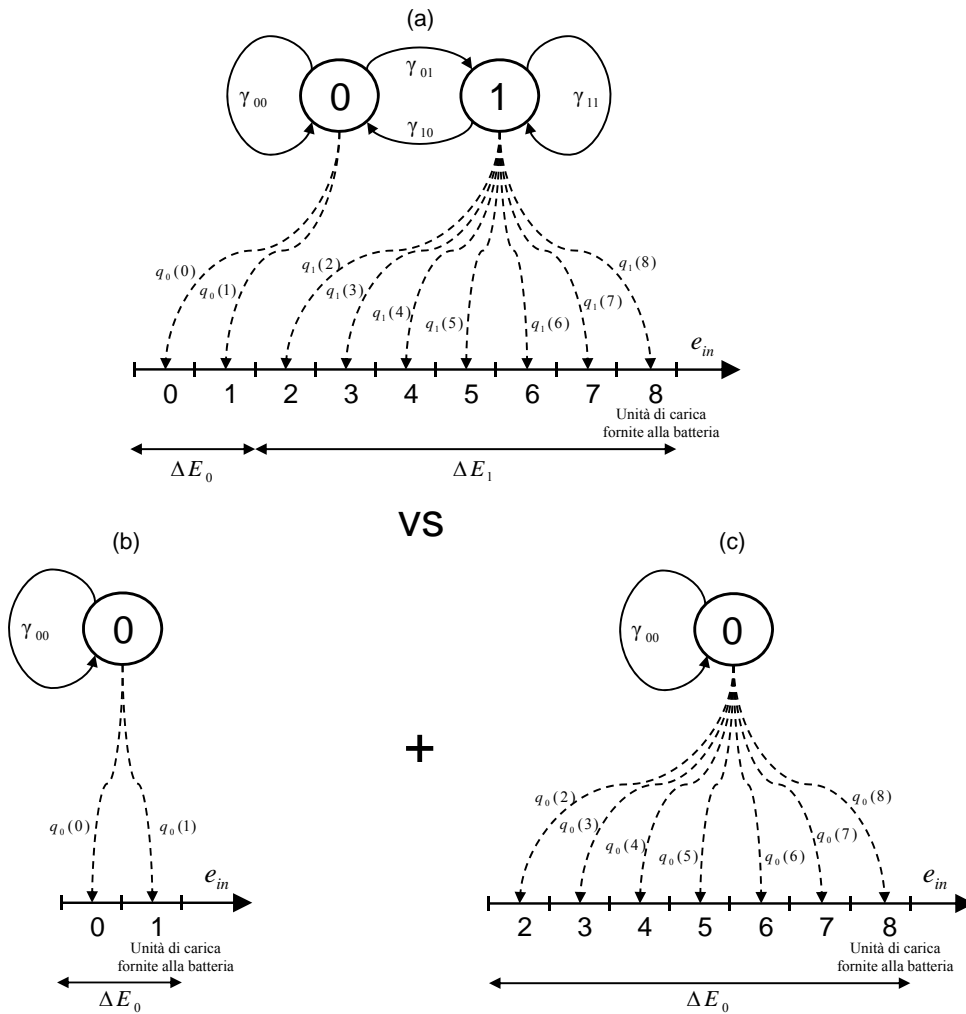


Figura 5.12 (a) Modello a due stati: $x = 0$ per la fase notturna, $x = 1$ per la fase diurna. (b) Modello a stato singolo per la fase notturna. (c) Modello a stato singolo per la fase diurna.

rappresentare la fonte rinnovabile durante il dì è possibile utilizzare il modello a stato singolo di figura 5.12(c), con $\Delta E_0 = \{2, 3, \dots, 8\}$ e $q_0(\cdot) = q'(\cdot)$ uniforme sull'insieme di valori $\{2, 3, \dots, 8\}$. Adottando questo modello, la corrispondente politica ottima μ^1 ottenuta sarà adeguata per le ore diurne. In questo modo, si può pensare di ricavare la politica ottima complessiva per il sensore radio combinando tra loro μ^0 e μ^1 , utilizzando la prima durante le fasi notturne e la seconda durante il dì.

A questo punto si intende confrontare la politica ottima μ^* , ottenuta adottando il modello a due stati, con quella ricavata dalla combinazione di μ^0 e μ^1 , calcolate utilizzando i due distinti modelli a stato singolo. In particolare, la politica μ^0 dovrà

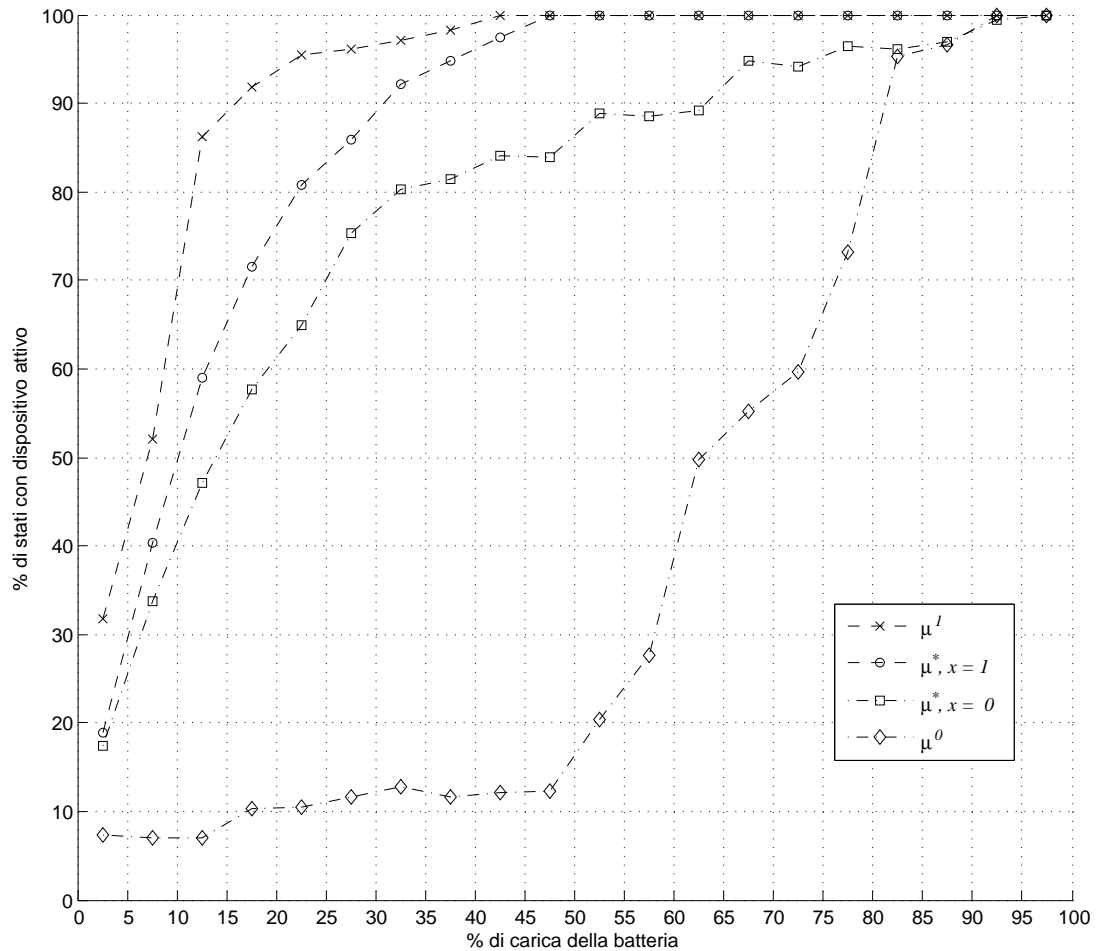


Figura 5.13 Livello di attività del sensore radio in funzione dello stato di carica della batteria, secondo le diverse politiche confrontate.

essere comparata con la politica μ^* , considerando per quest'ultima solo gli stati in cui $x = 0$, mentre μ^1 sarà confrontata con μ^* valutata esclusivamente negli stati con $x = 1$. Valutare, per μ^* , solamente gli stati in cui $x = 0$ significa, infatti, considerare tale politica limitatamente alle ore notturne, mentre analizzare gli stati con $x = 1$ consente di concentrare l'attenzione solo sulle ore diurne. Il confronto viene fatto valutando, ancora una volta, il livello di attività del sensore radio, secondo le diverse politiche, in funzione dello stato di carica della batteria. I risultati ottenuti sono riportati in figura 5.13.

L'aspetto più importante, che è possibile osservare immediatamente, è che le politiche ottenute per la fase diurna e per quella notturna, adottando prima il modello a due

stati e poi quello a stato singolo, sono tra loro *diverse*. Nello specifico, la politica μ^* nella fase diurna porta il sensore radio ad assumere, in generale, un comportamento più prudente per quanto riguarda il consumo energetico rispetto a quello specificato dalla politica μ^1 ; ciò è evidente soprattutto quando il livello di carica della batteria è basso. Viceversa, nelle ore notturne la politica μ^* induce il dispositivo ad essere molto più aggressivo rispetto a quanto indicato dalla politica μ^0 .

Tale discrepanza è dovuta proprio ai diversi modelli adottati per la determinazione delle politiche confrontate. Le azioni ottime specificate da μ^* per la fase diurna sono calcolate tenendo conto che la sorgente di energia alterna tale fase con quella notturna e che quindi per un certo periodo di tempo la fonte non sarà in grado di ricaricare la batteria. Questa consapevolezza non c'è, invece, nel calcolo della politica μ^1 la quale è determinata assumendo di rappresentare la sorgente come se esistesse esclusivamente la fase diurna. Ecco che allora è ragionevole che il comportamento ottimo per le ore diurne specificato da μ^* sia più prudente rispetto a quello indicato da μ^1 : il sensore radio deve essere più parsimonioso durante il giorno in previsione delle ore notturne durante le quali l'apporto energetico della fonte rinnovabile sarà molto scarso. Per quanto riguarda il confronto tra μ^* e μ^0 , il discorso è analogo: secondo la politica μ^* il dispositivo può essere più attivo di notte, rispetto a quanto indicato da μ^0 , in previsione della fase diurna durante la quale la sorgente di energia sarà in grado di ricaricare la batteria in modo più significativo.

In conclusione, i risultati appena discussi evidenziano quanto sia importante adottare il modello a due stati per rappresentare una fonte rinnovabile che alterna fasi in cui il contributo energetico è minimo con altre in cui risulta più significativo. In questo modo, infatti, il calcolo della politica ottima è *consapevole* di quello che è il particolare comportamento della sorgente di energia. Tale politica ottima non può essere ottenuta, invece, considerando separatamente le due fasi ed applicando la procedura di ottimizzazione per ciascuna di esse. Ciò che è importante modellare, dunque, è la *correlazione temporale* esistente tra le due fasi che caratterizzano il comportamento della fonte: mentre il modello a due stati consente di catturare ed includere nell'ottimizzazione tale correlazione, il modello a stato singolo non permette di rappresentarla e di tenerne

conto nel calcolo della politica ottima.

5.3.6 Azioni specificate dalla politica ottima

Nei risultati visti fino a questo momento l'analisi delle politiche ottime è stata condotta considerando per i diversi livelli di carica della batteria la percentuale di stati con dispositivo attivo. In questa sezione, invece, l'intenzione è quella di valutare, in funzione dello stato dei buffer e della batteria, la percentuale di stati in cui la politica ottima indica che il sensore radio debba compiere una certa azione $a \in \mathcal{A}$.

Le ipotesi operative considerate sono esattamente le stesse poste in sezione 5.3.4, con l'unica differenza che in questo caso la radio e il sensore vengono assunti mutuamente esclusivi. I risultati relativi alla politica ottima ottenuta con queste ipotesi sono illustrati in figura 5.14.

Si considerano separatamente le quattro possibili configurazioni dei buffer, ricordando che per ipotesi la capacità di ciascun buffer è di un solo pacchetto. In ognuno di questi casi, che rappresentano distinte situazioni in cui il dispositivo può trovarsi ad operare, le azioni ammissibili sono diverse. Naturalmente l'azione (ID) , che prevede che il sensore radio non esegua alcuna operazione, è sempre possibile.

Si valuta, innanzitutto, il caso in cui entrambi i buffer siano vuoti. In questa situazione, il dispositivo, oltre a rimanere a riposo (ID) , può eseguire il sensing (SX) oppure accendere la radio in ricezione (RX) . Come si osserva in figura 5.14(a), per bassi livelli di carica della batteria l'azione ottima predominante è (ID) : il sensore radio deve utilizzare in modo parsimonioso la poca energia a disposizione in modo da facilitare il processo di ricarica. All'aumentare del livello di carica della batteria, invece, il sistema è portato ad essere sempre più attivo e le azioni ottime prevalenti diventano (SX) e (RX) . In particolare, si nota una situazione di assoluto equilibrio tra queste due azioni, dovuta alle condizioni di equità imposte tra l'attività di rilevazione e quella di inoltro che portano il sensore radio a dedicarsi in egual misura al sensing e alla ricezione di un nuovo pacchetto.

Dal punto di vista qualitativo la situazione è analoga nel caso in cui i buffer siano entrambi pieni, come testimonia la figura 5.14(d). Anche in questa circostanza, quando

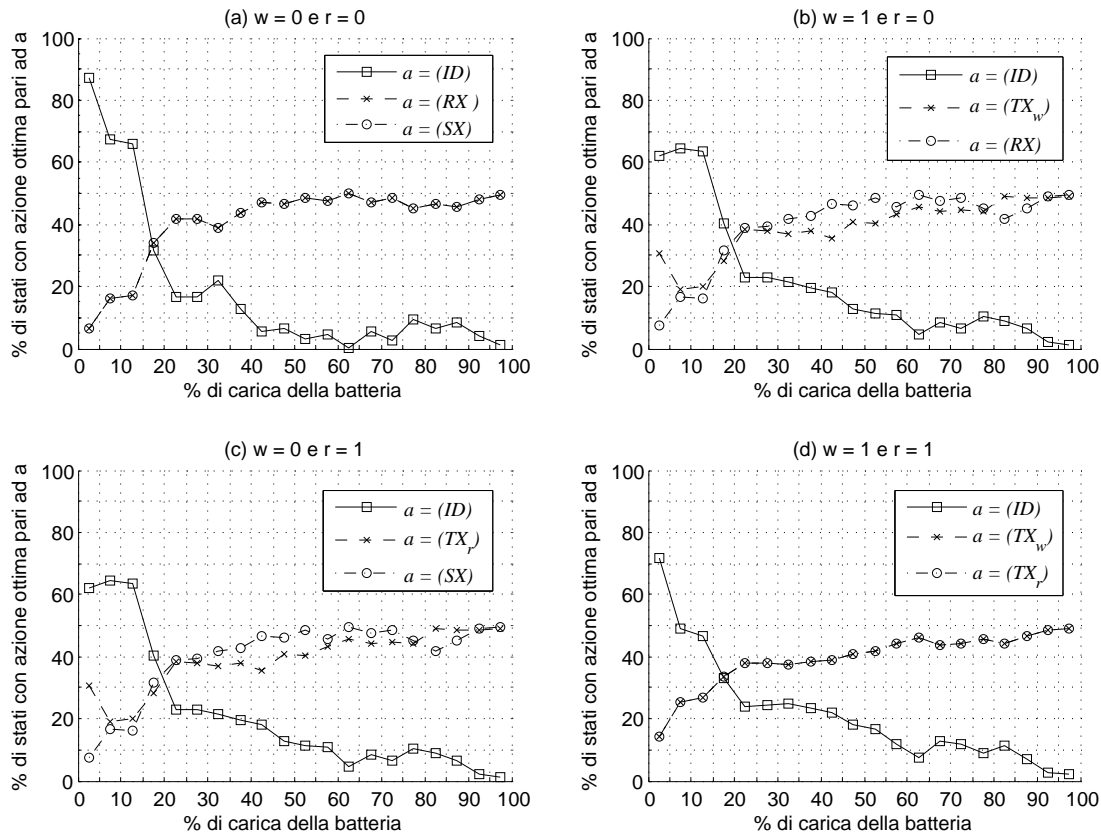


Figura 5.14 Azioni ottime in funzione dello stato dei buffer e della batteria. (a) Entrambi i buffer vuoti. (b) Buffer per il traffico locale pieno e l'altro vuoto. (c) Buffer per il traffico esogeno pieno e l'altro vuoto. (d) Entrambi i buffer pieni.

la batteria è molto scarica l'azione ottima prevalente è (ID) , mentre per livelli di carica superiori il dispositivo è maggiormente attivo e si dedica allo stesso modo alle altre due azioni ammissibili (TX_w) e (TX_r) .

I risultati non sono molto diversi nemmeno quando uno dei due buffer è pieno e l'altro è vuoto, come evidenziano la figura 5.14(b) e la figura 5.14(c). Nel caso sia pieno il buffer per il traffico locale, il sensore radio può decidere se trasmettere il pacchetto in esso contenuto (TX_w) oppure ricevere un pacchetto esogeno (RX) . Anche in questa situazione, le condizioni di equità tra l'attività di inoltra e quella di rilevazione spingono il dispositivo a dare sostanzialmente la stessa rilevanza alle due azioni. Esattamente speculare è la situazione nel caso sia il buffer per il traffico esogeno ad essere pieno: le azioni (TX_r) e (SX) sostituiscono rispettivamente (TX_w) e (RX) .

5.4 Riepilogo dei risultati ottenuti

I risultati discussi in questo capitolo mettono in evidenza alcuni aspetti interessanti del modello proposto per il sensore radio e delle politiche ottime determinate utilizzando l'ottimizzatore.

In particolare, dai risultati ottenuti mediante simulazioni si evince che:

- Una volta fissate le condizioni operative in cui si trova il sensore radio, la politica di scheduling determinata utilizzando l'ottimizzatore è quella che porta il dispositivo ad ottenere le prestazioni migliori. Questo testimonia che il modello proposto è ben definito e che la procedura di ottimizzazione implementata funziona correttamente.
- Variando il parametro β del modello si ottengono gli effetti desiderati sulle prestazioni del sensore radio. Tale parametro può dunque essere opportunamente regolato per spingere il dispositivo a preferire una tra le due attività di inoltro e rilevazione.
- Le prestazioni del sensore radio, raggiunte seguendo la politica ottima, rispondono nel modo atteso alle variazioni del consumo energetico dell'attività di sensing. Nello specifico, quando il sensing di un nuovo pacchetto implica un dispendio di energia minore rispetto a quello richiesto dalla ricezione, il dispositivo predilige l'attività di rilevazione. La situazione è invertita nel caso contrario.
- All'aumentare della quantità di energia massima che la fonte rinnovabile è in grado di fornire alla batteria in ciascuno slot, sono progressivamente migliori le prestazioni del sensore radio ottenute in accordo con la politica ottima.

I risultati ottenuti mediante analisi delle politiche ottime, invece, evidenziano che:

- Il grado di incertezza che si ha sul contributo energetico fornito in ogni slot dalla fonte rinnovabile influisce sul livello di attività del sensore radio soprattutto quando la batteria è scarica: se l'incertezza è bassa, la politica ottima spinge il dispositivo ad essere più attivo rispetto a quando l'incertezza è maggiore.

- Variando il livello di guardia della batteria, l'effetto ottenuto sulla corrispondente politica ottima è quello desiderato: il sensore radio è indotto a rimanere a riposo quando il livello di carica della batteria scende al di sotto della soglia prefissata.
- Se il consumo energetico dell'attività di sensing è inferiore a quello relativo alla ricezione, la politica ottima spinge il sensore radio ad essere più attivo negli slot in cui $v = 0$ piuttosto che in quelli in cui $v = 1$. Ciò è dovuto al fatto che l'attività di inoltro, in queste ipotesi, è complessivamente più onerosa dal punto di vista energetico rispetto a quella di rilevazione.
- Quando la fonte di energia rinnovabile ha un comportamento a due fasi, la politica ottima porta il sensore radio ad essere maggiormente attivo negli stati in cui $x = 1$ rispetto a quelli con $x = 0$. Nel primo caso, infatti, il dispositivo può contare su un apporto energetico non nullo da parte della sorgente, mentre nel secondo caso il contributo è nullo.
- Rappresentare una fonte di energia rinnovabile a due fasi mediante il modello a due stati fa sì che il calcolo della politica ottima sia consapevole del particolare comportamento della sorgente. Diversa è la politica ottenuta considerando separatamente le due fasi ed eseguendo l'ottimizzazione su ciascuna di esse. Mentre con il modello a due stati, infatti, è possibile modellare la correlazione temporale esistente tra le due fasi che caratterizzano la fonte, con il modello a stato singolo ciò non è possibile.

6

Conclusioni e sviluppi futuri

Il lavoro svolto e presentato in questa tesi si propone come valido punto di partenza per lo studio di una realtà complessa come quella dei sensori radio alimentati da fonti di energia rinnovabili. La possibilità di utilizzare l'energia presente nell'ambiente circostante per ricaricare la batteria dei nodi di una WSN sembra essere una soluzione molto interessante ed efficace per risolvere il problema dell'autonomia dei dispositivi della rete, soprattutto laddove l'intervento umano per cambiare o ricaricare manualmente la batteria è precluso o eccessivamente oneroso. Sfruttare fonti di energia rinnovabili, il cui comportamento non è per loro natura prevedibile deterministicamente, pone però degli importanti interrogativi su quali possano essere le prestazioni raggiungibili compatibilmente con la quantità di energia a disposizione e con le modalità con cui essa viene fornita al sensore radio. Interessa capire, poi, quale sia effettivamente il comportamento ottimo, ovvero la politica di scheduling ottima, che consenta al dispositivo di utilizzare al meglio tali sorgenti di energia e ottenere, dunque, le prestazioni massime.

In tale contesto, il principale contributo di questa tesi è il modello proposto per un generico sensore radio alimentato da una fonte di energia rinnovabile. Tale modello risponde a due importanti esigenze: giungere ad una rappresentazione formale del sistema considerato e disporre di uno strumento matematico utile per il calcolo delle politiche ottime.

Nella fase di realizzazione del modello si è puntato ad ottenere una rappresentazione del sistema il più possibile *generale*. Di fatto, il modello proposto è applicabile indipendentemente dalla particolare fonte di energia rinnovabile considerata e dalle specifiche caratteristiche del sensore radio. Si sono cercati di catturare, infatti, gli aspetti fondamentali che contraddistinguono un dispositivo di questo tipo. Sono stati considerati gli elementi costitutivi tipici come la radio, il sensore, la batteria e si è fatto riferimento alle due attività principali generalmente svolte da un sensore radio: la rilevazione e l'inoltro.

Il modello, poi, è stato costruito con l'intento di descrivere il funzionamento del sistema in modo *formale e rigoroso*. A tal fine si è fatto riferimento alla programmazione dinamica stocastica e si sono utilizzate le definizioni e gli strumenti forniti dalla teoria dei processi decisionali di Markov. In particolare, rappresentare il sensore radio mediante un processo decisionale di Markov ad orizzonte infinito ha permesso di formulare in veste matematica il problema di determinare la politica di scheduling ottima che consenta al dispositivo di raggiungere le prestazioni massime nelle condizioni in cui si trova ad operare. Tale problema, infatti, è stato ricondotto al calcolo della politica ottima secondo il criterio del costo medio per slot per il processo decisionale rappresentativo del sensore radio.

Un'ulteriore caratteristica molto interessante del modello realizzato è quella di essere altamente *parametrizzabile*. Sono stati definiti una serie di parametri sui quali è possibile intervenire e se ne sono discussi gli effetti sul modello. Impostando in modo opportuno i valori di tali parametri è possibile definire le particolari condizioni operative in cui si trova il sensore radio considerato. Questo aspetto è molto importante e contribuisce a rendere il modello generale, potendolo adattare alle diverse situazioni contingenti semplicemente agendo sui parametri.

Una volta costruito il modello per il sensore radio, l'attenzione è stata rivolta all'individuazione di una procedura di ottimizzazione valida per il calcolo delle politiche ottime. Nello specifico, analizzando le caratteristiche strutturali del processo decisionale rappresentativo del dispositivo, si è giunti a dimostrare l'applicabilità di una particolare versione dell'algoritmo noto in letteratura come *value iteration*.

Nella definizione del processo decisionale di Markov rappresentativo del sensore radio si è mirato a mantenere il più possibile contenuta la complessità, così da ottenere un modello trattabile dal punto di vista computazionale. Ciò ha consentito l'implementazione in C++ del modello stesso e del relativo algoritmo di ottimizzazione. Nello specifico, si è realizzato un *ottimizzatore* in grado di calcolare, in base ai valori stabiliti per i parametri, la corrispondente politica di scheduling ottima per il sensore radio. L'ottimizzatore così ottenuto è stato ampiamente sfruttato, come utile strumento di calcolo, nella fase dedicata alle simulazioni ed analisi.

Prima di tutto è stata *simulata* l'evoluzione del sensore radio in accordo con le politiche ottime calcolate dall'ottimizzatore al variare dei parametri. I risultati ottenuti hanno evidenziato l'effettiva validità dell'algoritmo di ottimizzazione e messo in luce il ruolo dei vari parametri del modello.

Successivamente, a conclusione del lavoro, si è condotta una prima *analisi* qualitativa delle politiche ottime ottenute al variare delle condizioni operative. Sono state individuate alcune caratteristiche significative di tali politiche, soprattutto in relazione alla descrizione statistica della fonte di energia rinnovabile. In particolare, si è sottolineata l'importanza del modello a due stati per rappresentare una sorgente con un comportamento a due fasi.

Questa tesi, nel suo complesso, ha il merito di gettare le basi teorico matematiche necessarie per studiare in modo analitico e sistematico quale siano le possibilità offerte, dal punto di vista applicativo, dall'utilizzo di fonti di energia rinnovabili per l'alimentazione dei nodi di una WSN. Nel lavoro svolto, che rappresenta un primo passo in tale direzione, possono essere messi in evidenza alcuni aspetti che meritano in futuro un ulteriore approfondimento.

La necessità di ottenere un modello per il sensore radio il più generale possibile ha portato inevitabilmente a lavorare ad un alto livello di astrazione. Il vantaggio di questo approccio è quello di poter adattare il modello realizzato a diverse situazioni operative reali. Nello specifico, una volta stabilito il particolare dispositivo radio considerato e la fonte di energia rinnovabile da adottare per alimentarlo, è sufficiente fissare opportunamente i parametri del modello per descrivere formalmente il sistema. Allora, sarà utile studiare in modo più dettagliato, mediante misure sul campo, come tradurre le specifiche tecniche del particolare sensore radio considerato in valori adeguati per i parametri del modello, così che quest'ultimo sia in grado di rappresentare il sistema nel modo più accurato possibile.

Stabilita la particolare fonte di energia rinnovabile da utilizzare, può essere utile, poi, cercare di ottenere per essa una caratterizzazione statistica più accurata e specifica rispetto a quella generale discussa in questa tesi. Si consideri ad esempio l'energia solare: misurando ed analizzando alcuni profili energetici che specificano la quantità di

energia che la sorgente è in grado di fornire alla batteria nel corso del tempo, è possibile ricavare un modello più preciso. In particolare, anziché utilizzare il semplice modello a due stati, può essere interessante rappresentare la fonte di energia mediante il modello generale a più stati. Non solo, anche le distribuzioni $q_x(\cdot)$, che specificano il numero di unità di carica che alimentano la batteria in funzione dello stato x occupato dalla sorgente, possono essere definite in modo più accurato.

Inoltre, può essere interessante introdurre nel modello una descrizione più precisa anche per quanto riguarda la batteria. Si può pensare di rappresentare alcuni fenomeni fisici comuni alle batterie tipicamente utilizzate nelle applicazioni reali. Si può tenere conto, ad esempio, del cosiddetto *rate capacity effect* [12], ovvero della dipendenza della capacità totale della batteria dall'intensità del flusso che la scarica. Un altro esempio è il *recovery effect* [12]: nei momenti in cui non è sollecitata, la batteria è in grado di recuperare autonomamente un certa quantità di energia che dipende dallo stato di carica in cui si trova e dalla durata dell'intervallo di pausa. Descrivere anche questi fenomeni richiede, tuttavia, un modello più complesso per la batteria e, in particolare, un numero maggiore di stati necessari per rappresentarne il funzionamento. Ciò si scontra, naturalmente, con l'esigenza di mantenere il più possibile contenuta la complessità del processo decisionale rappresentativo del sensore radio. Inoltre, tali fenomeni devono essere rappresentati in modo corretto per non introdurre degli effetti spuri e incontrollabili nei risultati.

Un altro aspetto su cui sarà particolarmente interessante soffermarsi è lo studio delle politiche di scheduling ottime ottenute dall'ottimizzatore al variare delle condizioni operative. Sarà molto utile condurre un'analisi approfondita ed analitica finalizzata ad estrapolare da tali politiche delle regole di comportamento generali per il sensore radio. L'idea è quella di ottenere, una volta note le condizioni operative del sistema, delle indicazioni di massima su quale sia il comportamento ottimo per il dispositivo senza dover calcolare esplicitamente la politica ottima mediante l'ottimizzatore.

Infine, si può pensare di estendere il modello proposto in questa tesi, per un singolo sensore radio, a tutti i nodi di una WSN. L'obiettivo finale è quello di realizzare un modello generale valido per un'intera rete composta da sensori radio alimentati da una

fonte di energia rinnovabile. Così facendo si avrebbe a disposizione uno strumento molto utile per ottimizzare protocolli, come quelli di routing, accesso al mezzo e disseminazione per reti di questo tipo.

Bibliografia

- [1] D. P. Bertsekas, *Dynamic Programming and Optimal Control: vol. I*, 3rd ed., Athena Scientific, 2007.
- [2] D. P. Bertsekas, *Dynamic Programming and Optimal Control: vol. II*, 3rd ed., Athena Scientific, 2007.
- [3] D. P. Bertsekas, J. N. Tsitsiklis, *Introduction to Probability*, Athena Scientific, 2002.
- [4] D. Blackwell, "Discrete Dynamic Programming," *Ann. Math. Statist.*, vol. 33, pp. 719-726, 1962.
- [5] G. Cariolaro, G. Pierobon, *Processi Aleatori*, 4th ed., Libreria Progetto, 2008.
- [6] K. Chebrolu, B. Raman, N. Mishra, P.K. Valiveti, R. Kumar, "Brimon: A Sensor Network System for Railway Bridge Monitoring," *Proc. 6th ACM International Conference on Mobile Systems, Applications, and Services*, pp. 2-14, 2008.
- [7] P. Desnoyers, D. Ganesan, H. Li, M. Li, P. Shenoy, "PRESTO: A Predictive Storage Architecture for Sensor Networks," *Proc. 10th USENIX Conference on Hot Topics in Operating Systems*, p. 23, 2005.
- [8] B.L. Fox, D.M. Landi, "An algorithm for identifying the ergodic subchains and transient states of a stochastic matrix," *Communications of the ACM*, vol. 11, no. 9, pp. 619-621, 1968.
- [9] W. Heinzelman, A. Chandrakasan, H. Balakrishnan, "Energy-efficient Communication Protocol for Wireless MicroSensor Networks," *Proc. 33rd Annual Hawaii International Conference on System Sciences*, p. 8020, vol. 8, 2000.

-
- [10] C. Intanagonwiwat, R. Govindan, D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," *Proc. 6th ACM Annual International Conference on Mobile Computing and Networking*, pp. 56-67, 2000.
- [11] X. Jiang, J. Polastre, D. Culler, "Perpetual Environmentally Powered Sensor Networks," *Fourth International Symposium on Information Processing in Sensor Networks*, pp. 463-468, 2005.
- [12] M.R. Jongerden, B.R. Haverkort, "Which battery model to use?," *IET Software*, vol. 3, no. 6, pp. 445-457, 2010.
- [13] S. Karlin, H. M. Taylor, *An Introduction to Stochastic Modeling*, 3rd ed., Academic Press, 1998.
- [14] M. Karpiriski, A. Senart, V. Cahill, "Sensor Networks for Smart Roads," *Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 310-314, 2006.
- [15] S. Kianpisheh, N.M. Charkari, "Dynamic Power Management for Sensor Node in WSN Using Average Reward MDP," *Proceedings of the 4th International Conference on Wireless Algorithms, Systems, and Applications*, pp. 53-61, 2009.
- [16] S. Kumar, T. H. Lai, J. Balogh, "On k-coverage in a Mostly Sleeping Sensor Network," *Wireless Networks*, vol. 14, no. 3, pp. 277-294, 2008.
- [17] R.G. Lee, K.C. Chen, S.S. Chiang, C.C. Lai, H.S. Liu, M.S. Wei, "A Backup Routing with Wireless Sensor Network for Bridge Monitoring System," *Annual IEEE Conference on Communication Networks and Services Research*, pp. 157-161, 2006.
- [18] H. Liu, A. Chandra, J. Srivastava, "eSENSE: Energy Efficient Stochastic Sensing Framework for Wireless Sensor Platforms," *Proc. Fifth International Conference on Information Processing in Sensor Networks*, pp. 235-242, 2006.
- [19] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, J. Anderson, "Wireless Sensor Networks for Habitat Monitoring," *Proc. 1st ACM International Workshop on Wireless Sensor Networks and Applications*, pp. 88-97, 2002.

- [20] D. Niyato, E. Hossain, A. Fallahi, "Sleep and Wakeup Strategies in Solar-Powered Wireless Sensor/Mesh Networks: Performance Analysis and Optimization," *IEEE Trans. Mobile Comput.*, vol. 6, no. 2, pp. 221-236, 2007.
- [21] M. L. Puterman, *Markov Decision Process: Discrete Stochastic Dynamic Programming*, Wiley-Interscience, 2005.
- [22] Q. Qiu, Q. Wu, M. Pedram, "Stochastic Modeling of a Power-Managed System-Construction and Optimization," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 20, no. 10, pp. 1200-1217, 2001.
- [23] V. Raghunathan, A. Kansal, J. Hsu, J. Friedman, M. Srivastava, "Design Considerations for Solar Energy Harvesting Wireless Embedded Systems," *Fourth International Symposium on Information Processing in Sensor Networks*, pp. 457-462, 2005.
- [24] P. Rong, M. Pedram, "Battery-Aware Power Management Based on Markovian Decision Processes," *IEEE Trans. on Computer Aided Design*, Vol. 25, No. 7, 2006.
- [25] L. I. Sennot, *Stochastic Dynamic Programming and the Control of Queueing System*, Wiley-Interscience, 1998.
- [26] S. Sudevalayam, P. Kulkarni, "Energy Harvesting Sensor Nodes: Survey and Implications," *IEEE Communications Surveys and Tutorials*, pp. 1-19, 2010.
- [27] A.E. Susu, A. Acquaviva, D. Atienza, G.D. Micheli, "Stochastic Modeling and Analysis for Environmentally Powered Wireless Sensor Nodes," *Proc. of the IEEE Intl Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, pp. 11-20, 2008.
- [28] J. Taneja, J. Jeong, D. Culler, "Design, Modeling, and Capacity Planning for Micro-solar Power Sensor Networks," *Proc. 7th International Conference on Information Processing in Sensor Networks*, pp. 407-418, 2008.
- [29] Texas Instruments, "Chipcon CC2420: 2.4 GHz IEEE 802.15.4 / ZigBee-ready RF Transceiver," <http://focus.ti.com/lit/ds/symlink/cc2420.pdf>, 2007.

-
- [30] L. Wang, Y. Xiao, "A Survey of energy-efficient scheduling mechanisms in sensor networks," *Mobile Networks and Applications*, vol. 11, pp. 723- 740, 2006.
- [31] G. Werner-Allen, K. Lorincz, M. Ruiz, O. Marcillo, J. Johnson, J. Lees, M. Welsh, "Deploying a Wireless Sensor Network on an Active Volcano," *IEEE Internet Comput.*, vol. 10, no. 2, pp. 18-25, 2006.
- [32] W. Ye, J. Heidemann, D. Estrin, "An Energy-efficient MAC Protocol for Wireless Sensor Networks," *INFOCOM 2002. Proc. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, pp. 1567-1576, 2002.