



UNIVERSITÀ DI FIRENZE
DIPARTIMENTO DI INFORMATICA

Corso di Laurea Triennale in Ingegneria Informatica

*Controllo di copertura dinamico per il
monitoraggio cooperativo in tempo reale*

Relatore

Giorgio Battistelli

Correlatore

Nicola Forti

Candidato

Ivan Necerini

Anno accademico 2023 / 2024

Indice

1	Introduzione	2
1.1	Descrizione del problema	4
1.2	Formulazione matematica del problema	5
1.3	Soluzioni proposte	6
2	Algoritmi	8
2.1	Algoritmo di coverage V1: algoritmo di ascesa del gradiente classico .	9
2.2	Algoritmo V2: aggiunta di un moto Browniano	10
2.3	Algoritmo V3: aggiunta di un potenziale repulsivo	13
3	Implementazione degli algoritmi	14
3.1	Descrizione del codice (e tabelle con pseudocodice)	14
3.2	Librerie Python utilizzate	14
4	Simulazione e discussione dei risultati	15
5	Conclusioni	16

1 Introduzione

Il controllo di copertura dinamico è un problema centrale in molti campi applicativi, tra cui la robotica, la sorveglianza, la gestione delle emergenze e le operazioni militari. L'obiettivo principale di questo problema è quello di garantire che un'area di interesse sia monitorata in modo efficace e continuo da una flotta di agenti mobili, come droni o veicoli autonomi, equipaggiati con sensori. *Questo compito diventa particolarmente complesso e, interessante e pratico quando l'ambiente è dinamico e i punti di interesse (targets) sono in movimento.* Il controllo dinamico della copertura è dunque un tipo di controllo cooperativo che richiede a un sistema multi-agente di monitorare dinamicamente un'area di interesse nel tempo (2). Questo tipo di controllo è ampiamente utilizzato in una varietà di applicazioni, come esplorazione, ricognizione, sorveglianza e tracciamento dei bersagli.

Nella letteratura, esistono due principali categorie di controllo della copertura: statica e dinamica. Il controllo della copertura statica si occupa di determinare il posizionamento ottimale degli agenti in modo che le aree di interesse siano completamente coperte dall'unione delle zone di rilevamento degli agenti. Al contrario, il controllo della copertura dinamica sfrutta la mobilità degli agenti per monitorare continuamente le aree di interesse nel tempo. In altre parole, ogni area deve essere monitorata da agenti mobili per un periodo sufficiente a garantire una copertura efficace. Il controllo della copertura statica, tuttavia, presuppone implicitamente la disponibilità di un numero sufficiente di agenti o la presenza di aree di interesse di dimensioni ridotte. In caso contrario, non vi è alcuna garanzia che tutte le aree possano essere completamente coperte. Queste limitazioni possono essere superate con il controllo della copertura dinamica, che permette il monitoraggio di vaste aree di interesse utilizzando un numero ridotto di agenti. Tuttavia, sia nel controllo della copertura statica che dinamica, è spesso necessario disporre di modelli dell'ambiente e della cinematica/dinamica degli agenti per sviluppare leggi di controllo del movimento efficaci (2).

Ottenere modelli accurati può risultare restrittivo e addirittura irrealizzabile in molte applicazioni pratiche. Per questo motivo, il presente lavoro è fortemente orientato a sviluppare un approccio privo di modelli, in cui gli agenti possano apprendere direttamente dalle interazioni con l'ambiente, ottenendo così un controllo dinamico della copertura.

Il Reinforcement Learning (RL) si presenta come una soluzione promettente per affrontare il problema del controllo di copertura dinamico. L'RL è un processo decisionale sequenziale in cui un agente interagisce con l'ambiente e apprende da esso, sviluppando politiche ottimali *senza la necessità di modelli predittivi espliciti.*

Tuttavia, l'RL tradizionale è principalmente limitato a spazi di azione discreti. Per affrontare problemi con spazi di azione continui o dati ad alta dimensione, il Deep Reinforcement Learning (DRL) integra reti neurali, offrendo soluzioni più efficaci. Ad esempio, i robot possono apprendere politiche di controllo direttamente dai dati delle telecamere piuttosto che da controller ingegnerizzati manualmente o da caratteristiche a bassa dimensione degli stati del robot. Recen-

temente, il focus della ricerca si è spostato verso il Multi-Agent Deep Reinforcement Learning (MADRL), poiché più agenti sono in grado di affrontare compiti complessi in modo più efficiente rispetto a un singolo agente. Il MADRL trova applicazioni in vari ambiti, tra cui giochi online multiplayer, robot cooperativi, controllo del traffico e sistemi di controllo dell'energia. Anche il controllo di copertura dinamico ha beneficiato dell'applicazione di tecniche MADRL, come dimostrato in vari studi.

Nel contesto di questo progetto, l'implementazione di algoritmi basati su DRL e MADRL rappresenta una possibile espansione futura. Questi approcci potrebbero migliorare ulteriormente le prestazioni del sistema multi-agente, consentendo agli agenti di apprendere politiche di controllo più sofisticate e coordinarsi in modo più efficace.

In questo lavoro, ci concentriamo sull'implementazione di un algoritmo di controllo di copertura basato sull'ascesa del gradiente (*Gradient Ascent Algorithm*), una tecnica di Reinforcement Learning (RL) che permette agli agenti di apprendere politiche di controllo ottimali attraverso l'interazione con l'ambiente. Questo approccio è particolarmente adatto per scenari in cui non si dispone di un modello completo dell'ambiente o dove le condizioni cambiano rapidamente nel tempo.

L'algoritmo di ascesa del gradiente è stato scelto per la sua capacità di ottimizzare direttamente una funzione obiettivo, in questo caso l'indice di copertura totale, che quantifica la qualità della copertura fornita dagli agenti sui targets. Questo metodo è stato esteso per considerare le limitazioni pratiche, come le capacità di rilevamento e comunicazione limitate degli agenti, nonché la necessità di mantenere una copertura sufficiente su tutti i targets durante l'intera durata della missione.

Il presente lavoro mira a sviluppare e implementare un algoritmo di controllo di copertura dinamico che sfrutti i principi del Reinforcement Learning e dell'ascesa del gradiente. Gli agenti mobili (droni) devono monitorare una serie di targets mobili, garantendo che ogni target sia coperto cumulativamente a un livello desiderato nel tempo. La formulazione matematica del problema, le soluzioni proposte e i dettagli dell'implementazione dell'algoritmo saranno discussi nelle sezioni successive.

1.1 Descrizione del problema

In questo lavoro, un piccolo gruppo di agenti mobili (*UAVs, droni*) con capacità di rilevamento limitate, posti ad un'altezza fissa, ha l'obiettivo di monitorare dinamicamente un'area popolata da un gran numero di punti di interesse (*Targets mobili*). Lo scenario di riferimento potrebbe essere, ad esempio, un'area portuale dove i targets sono rappresentati da imbarcazioni che si muovono continuamente nel tempo.

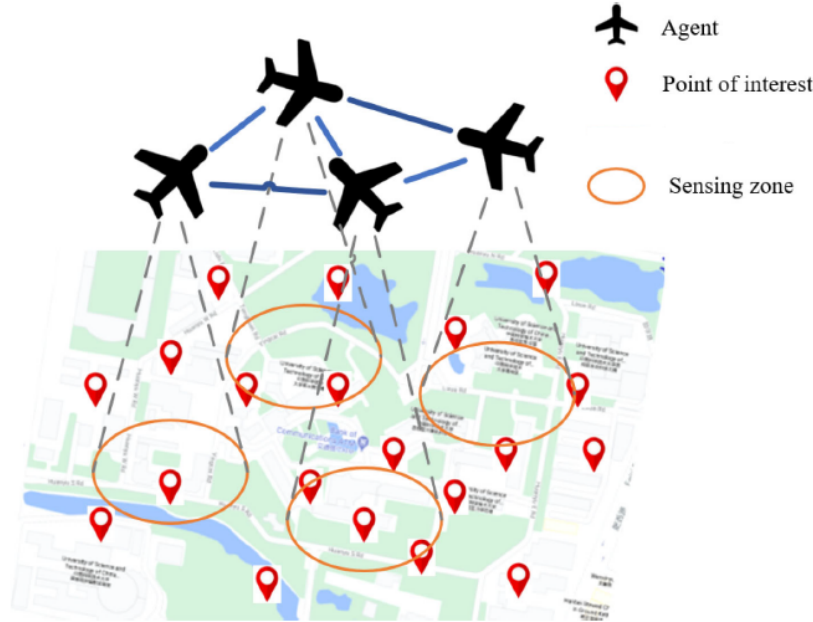


Figura 1: Possibile scenario di un controllo della copertura dinamica (2).

Il monitoraggio dinamico di un'area di interesse richiede che, istante per istante, i droni forniscano una visione accurata (o almeno sufficiente) di tutti i target mobili presenti. L'obiettivo generale può variare, come ad esempio l'evitare collisioni tra le imbarcazioni.

Per raggiungere questo obiettivo, è fondamentale garantire una *copertura sufficiente* dell'area. Nel contesto di questo esperimento, le traiettorie dei target saranno predefinite e verranno assegnate posizioni iniziali ai droni. **Il compito consiste dunque nello scrivere un algoritmo di copertura per determinare le traiettorie dei droni in modo tale che, al termine delle traiettorie dei target, venga raggiunto un grado sufficiente di copertura dell'area.** (vedi 2).

I droni dovranno quindi "seguire" i target, assicurando una buona visione dell'area di interesse per tutta la durata dell'esperimento. La definizione dettagliata del *grado di copertura sufficiente* verrà fornita nella sezione 1.3 della tesi.

1.2 Formulazione matematica del problema

Consideriamo M punti di interesse (**Targets dinamici**) distribuiti in uno spazio di lavoro bi-dimensionale W . Le posizioni dei targets sono denotate da $q_j(t) \in \mathbb{R}^2$, con $j \in \{1, \dots, M\}$. Un gruppo di N **Agenti mobili, dinamici** (*Agents*), ad un'altezza fissata, è incaricato di monitorare continuamente questi targets. Le posizioni degli agenti sono denotate da $p_i(t) \in \mathbb{R}^2$, con $i \in \{1, \dots, N\}$.

Si presume che ogni agente abbia una zona di rilevamento limitata, modellata da un'area discoidale con un certo *raggio di copertura* $r \in \mathbb{R}^+$ centrata su se stesso.

Poiché la qualità del rilevamento generalmente degrada con la distanza dal target, il rilevamento basato sulla distanza dell'agente i sul target j è caratterizzato da una **FUNZIONE DI MISURAZIONE**:

$$E_{ij}(p_i(t), q_j(t)) = \begin{cases} M_P \left(\frac{(l_{ij}(t)-r)^2}{r^4} \right), & \text{se } l_{ij}(t) \leq r, \\ 0, & \text{se } l_{ij}(t) > r, \end{cases} \quad (1)$$

dove $l_{ij}(t) = \|p_i(t) - q_j(t)\|$ è la distanza, all'istante t , tra il drone i e il target j , e $M_P \in \mathbb{R}^+$ è una qualità di rilevamento di picco. Il modello di rilevamento indica che la qualità del rilevamento del target j raggiunge il picco quando l'agente i coincide con esso, e diminuisce monotonamente quando l'agente i si allontana da esso (2).

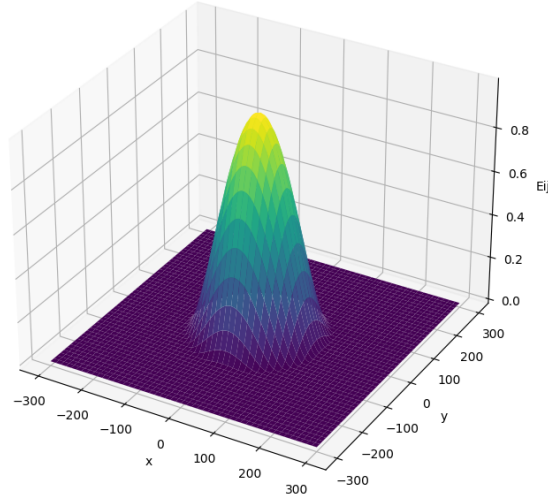


Figura 2: Funzione di misurazione, $mp=1$, $r=150$.

1.3 Soluzioni proposte

In questa sottosezione viene definita la procedura matematica che stabilisce il *grado di copertura sufficiente*, menzionato nel paragrafo 1.1.

Il controllo della copertura efficace basato su *RL* (Reinforcement Learning) in questo lavoro ha l'obiettivo di guidare gli agenti mobili affinché garantiscano che ogni target nella regione di interesse sia coperto cumulativamente a un livello desiderato nel tempo (2). Per ogni target j , viene definito un *indice di copertura* $E_j(t)$, che quantifica la qualità della copertura al tempo t fornita dai vari agenti. Ogni agente contribuisce all'indice di copertura in base alla propria vicinanza al target, secondo la seguente funzione di misurazione:

$$E_j(t) = \sum_{i=1}^N E_{ij}(p_i(t), q_j(t)), \quad j \text{ fisso} \quad (2)$$

Come indicato nella sezione 1.2, la funzione di misurazione $E_{ij}(p_i(t), q_j(t))$ varia nell'intervallo $[0, M_P]$. Di conseguenza, l'indice di copertura $E_j(t)$ apparterrà all'intervallo $[0, M_P \cdot N]$.

Successivamente, viene definito un *indice di copertura totale* $E(t)$, espresso come:

$$E(t) = \sum_{j=1}^M \sigma(E_j(t)) \quad (3)$$

dove:

$$\sigma(E_j(t)) = \frac{\tanh(E_j(t) - E^*) + 1}{2} \quad (4)$$

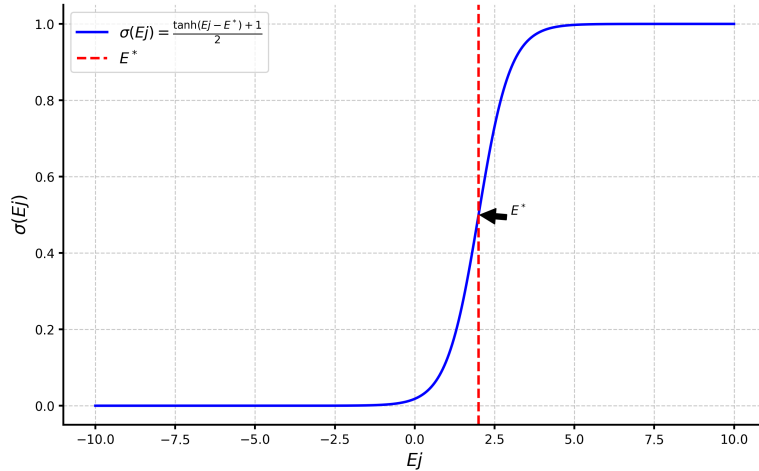


Figura 3: Funzione sigmoidale, $E^* = 2$

La definizione di queste funzioni e indici di copertura è fondamentale per raggiungere l'obiettivo del problema. La soglia E^* rappresenta un *lower bound index (LB)*, ovvero ogni indice di copertura di un target generico j al tempo t ($E_j(t)$) deve essere maggiore di questa soglia per considerare "accettabile" la copertura fornita dagli agenti. È importante notare che I PARAMETRI SONO FORTEMENTE INTERDIPENDENTI; infatti, in base a M_P , varia l'intervallo in cui è contenuta la funzione di misurazione $E_{ij}(p_i(t), q_j(t))$ e, di conseguenza, anche gli indici di copertura ($E_j(t)$). È quindi necessario settare i vari parametri caso per caso. Nella simulazione effettuata (vedi 4), sono stati fissati i seguenti parametri:

- $N = 4$. Rappresenta il numero degli agenti.
- $M = 10$. Rappresenta il numero dei target.
- $M_p = 1$. Definisce l'intervallo in cui è contenuta la funzione di misurazione e anche gli indici di copertura.
- $E^* = 2$. Definisce una soglia minima per gli indici di copertura.
- $duration = 150$. Durata in secondi delle traiettorie dei target.

Considerando questi parametri numerici, possiamo affermare che la funzione di misurazione varia tra 0 e 1, mentre gli indici di copertura variano tra 0 e 4. Impostare la soglia E^* a 2 implica che, nel complesso, desideriamo che almeno 2 agenti su 4 coprano ciascun target. Per raggiungere il valore 2, è possibile che 2 agenti siano esattamente sopra i target e abbiano quindi copertura massima, oppure che la somma delle funzioni di misurazione complessive sia pari a 2 (ogni agente copre in minima parte quel target, sommando le coperture si raggiunge la soglia).

La funzione sigmoideale è stata definita per tener conto di questo ragionamento sulla soglia. L'indice di copertura totale somma i singoli indici di copertura a cui è applicata la sigmoide, in modo che, durante il calcolo del gradiente (vedi 2), tutti i contributi degli agenti siano considerati.

Durante la simulazione, verranno generate delle posizioni iniziali degli agenti tali da avere una copertura leggermente inferiore al sufficiente nella situazione iniziale. L'obiettivo sarà quello di sviluppare un algoritmo di copertura dinamica (che verrà analizzato nella sezione 2) per creare le traiettorie degli agenti e indicare loro come muoversi, al fine di ottenere un *indice di copertura totale* finale significativamente maggiore rispetto a quello iniziale $E(duration - 1) > E(0)$.

Può succedere, come verrà analizzato in modo esaustivo nella sezione 4, che alcuni indici di copertura singoli al tempo finale $E_{ij}(duration - 1)$ siano azzerati o prossimi allo zero. Questo non rappresenta un problema, ma è un comportamento previsto dell'algoritmo di ascesa del gradiente (trattato nella sezione 2), che punta a massimizzare l'indice di copertura totale. Di conseguenza, alcuni target potrebbero non essere coperti nell'istante finale; tuttavia, ciò avvantaggia molti altri target che verranno coperti in maniera ottimale, contribuendo così ad aumentare l'indice totale.

2 Algoritmi

Come anticipato nella sezione 1, l'obiettivo di questo progetto è sviluppare un algoritmo di copertura dinamico (**dynamic coverage algorithm**) per determinare le traiettorie degli agenti, al fine di massimizzare l'indice di copertura totale E .

Gli algoritmi proposti si basano tutti sul **metodo di ascesa del gradiente** come fondamento implementativo. Un algoritmo basato sull'ascesa del gradiente è una tecnica di ottimizzazione iterativa utilizzata per individuare il massimo di una funzione obiettivo. L'idea principale consiste nell'aggiornare iterativamente le variabili nella direzione del gradiente della funzione obiettivo rispetto a queste variabili, al fine di aumentare il valore della funzione (4).

Nelle sezioni seguenti verranno illustrati tre differenti tipi di algoritmi di copertura dinamica basati sull'ascesa del gradiente. Le versioni seconda e terza saranno estese per risolvere problemi di collisione tra gli agenti nell'istante finale.

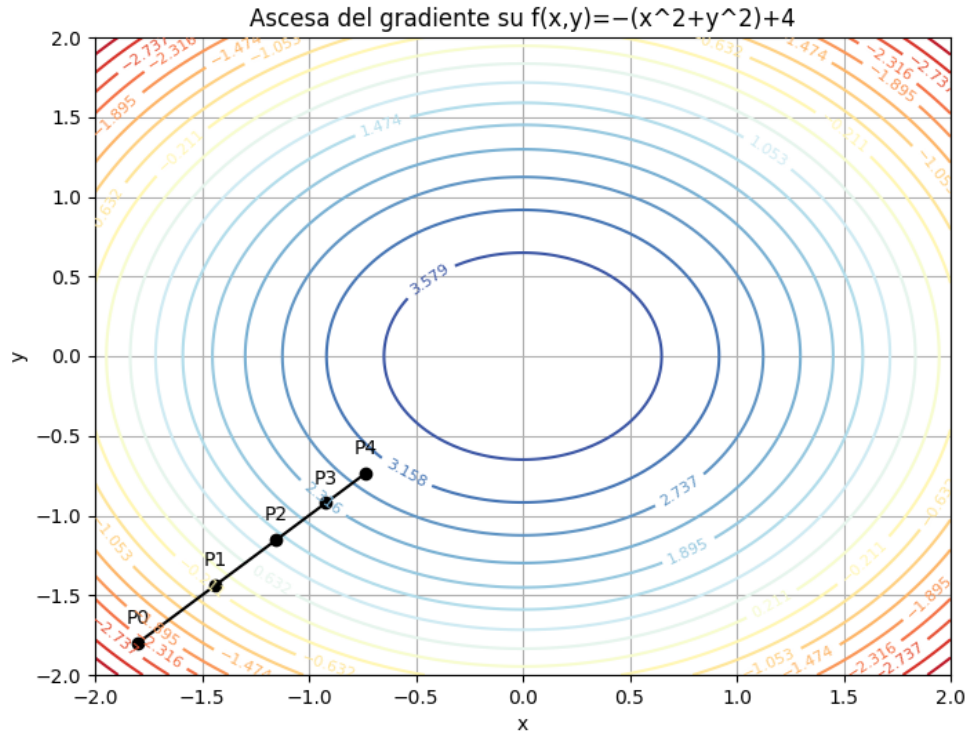


Figura 4: Ascesa del gradiente

2.1 Algoritmo di coverage V1: algoritmo di ascesa del gradiente classico

Nell'ambito del presente progetto di controllo della copertura dinamica, l'algoritmo di ascesa del gradiente viene impiegato per aggiornare le posizioni degli agenti mobili, come droni, determinando le loro traiettorie in modo da massimizzare una funzione obiettivo che rappresenta la copertura totale dell'area di interesse, ovvero l'indice di copertura totale $E(t)$. Gli agenti calcolano il gradiente della funzione di copertura rispetto alle loro posizioni e a quelle dei target al tempo corrente. Successivamente, aggiornano le loro traiettorie seguendo la direzione del gradiente, migliorando iterativamente la copertura complessiva dell'area.

Algorithm 1: V1: Algoritmo di coverage dinamico con ascesa del gradiente

Result: Massimizzare la copertura totale E

```

1 for ogni  $t$  do
2   for ogni agente  $i$  do
3     Decidere la direzione in cui muovere l'agente per massimizzare la copertura;
4     
$$\mu_i = \frac{\partial E}{\partial p_i} \Big|_{p_i=p_i(t), q_i=q_i(t)} (p_1, \dots, p_N, q_1, \dots, q_N)$$

5     L'agente si muove con passo di salita  $\epsilon$ ;
6      $p_i(t+1) = p_i(t) + \epsilon \mu_i(t);$ 
7   end
8 end
```

In questo contesto, ϵ rappresenta il passo di salita. Questo parametro è cruciale per l'aggiornamento delle posizioni degli agenti mobili ed è direttamente responsabile della velocità con cui gli agenti si muovono lungo la traiettoria determinata dal gradiente della funzione obiettivo.

Come descritto in dettaglio nella sezione 1.1, è possibile che al termine dell'esecuzione dell'algoritmo alcuni indici di copertura individuali $E_{ij}(\text{duration} - 1)$ risultino pari a zero o prossimi allo zero. Questo fenomeno non rappresenta un'anomalia, ma un comportamento previsto dell'algoritmo di ascesa del gradiente, il cui obiettivo principale è la massimizzazione dell'indice di copertura totale. Di conseguenza, alcuni target potrebbero non essere coperti al termine della simulazione.

Un problema emergente in questa situazione è che gli agenti, per massimizzare la copertura, possono convergere verso posizioni finali molto vicine tra loro al tempo $t = \text{duration}$, aumentando il rischio di collisione. Questo è un problema ben noto nel contesto del coverage dinamico, dove la priorità alla copertura totale può portare a densità elevate di agenti in aree specifiche.

Per mitigare il rischio di collisione e migliorare la distribuzione spaziale degli agenti, sono state sviluppate due ulteriori versioni dell'algoritmo. Queste versioni estendono la versione base dell'algoritmo di ascesa del gradiente, introducendo meccanismi che favoriscono una separazione adeguata tra gli agenti, garantendo una copertura efficace senza incorrere in collisioni esagerate.

Nel contesto del controllo di copertura dinamica, il problema della collisione tra agenti è un aspetto cruciale da gestire. Quando più agenti, come droni, operano nello stesso spazio per monitorare una serie di punti di interesse, c'è un rischio significativo che le loro traiettorie si intersechino, portando a collisioni. Questo è particolarmente problematico alla fine dell'esecuzione dell'algoritmo di copertura, quando gli agenti tendono a convergere verso le aree con alta densità di target.

Per affrontare questo problema, diverse strategie di controllo dell'evitamento delle collisioni (*collision avoidance*) sono state proposte in letteratura. Alcuni approcci utilizzano tecniche di ottimizzazione che includono vincoli di distanza minima tra gli agenti, mentre altri implementano algoritmi di controllo distribuito che permettono agli agenti di coordinarsi per evitare le collisioni in tempo reale. Un esempio di queste tecniche è l'uso di funzioni potenziali che generano forze repulsive tra gli agenti quando si avvicinano troppo l'uno all'altro (5) (come fatto in V3: sezione 2.3).

2.2 Algoritmo V2: aggiunta di un moto Browniano

In questa seconda versione dell'algoritmo, le posizioni degli agenti vengono aggiornate non solo utilizzando il gradiente della funzione obiettivo, ma anche includendo una componente stocastica derivante dal **moto browniano**. Questo approccio introduce una variazione casuale nelle traiettorie degli agenti, migliorando la loro capacità di esplorazione e riducendo il rischio di convergenza in posizioni subottimali.

Algorithm 2: V2: Algoritmo di coverage dinamico, aggiunta moto di Browniano

Result: Massimizzare la copertura totale E

```

1 for ogni  $t$  do
2   Calcolo della componente Browniana  $\eta_i(t)$ 
3   for ogni agente  $i$  do
4     Decidere la direzione in cui muovere l'agente per massimizzare la copertura;
5
6     
$$\mu_i = \frac{\partial E}{\partial p_i} \Big|_{p_i=p_i(t), q_i=q_i(t)} (p_1, \dots, p_N, q_1, \dots, q_N)$$

7     L'agente si muove con passo di salita  $\epsilon$ ;
8      $p_i(t+1) = p_i(t) + \epsilon \mu_i(t) + \eta_i(t);$ 
9   end
10 end
```

Il moto browniano, noto anche come movimento aleatorio, è un modello matematico utilizzato per descrivere il comportamento di particelle che si muovono in modo casuale in un fluido. Questo fenomeno prende il nome dal botanico Robert Brown, che per primo osservò il movimento irregolare di particelle microscopiche in sospensione nel 1827. Il moto browniano è caratterizzato da una traiettoria irregolare e imprevedibile, dove ogni passo successivo è indipendente dal precedente e segue una distribuzione di probabilità specifica (6).

Il moto browniano può essere modellato matematicamente come un processo stocastico, spesso descritto da equazioni differenziali stocastiche. In particolare, il modello di Wiener è comunemente utilizzato per rappresentare il moto browniano. Questo modello considera la traiettoria di una particella come una somma di incrementi casuali, ciascuno distribuito normalmente con media zero e varianza proporzionale al tempo. La funzione di densità di probabilità per la posizione della particella a un tempo futuro è quindi una distribuzione gaussiana, il cui centro si sposta con il tempo e la cui larghezza aumenta proporzionalmente alla radice quadrata del tempo (6).

Nel contesto degli algoritmi di ottimizzazione e controllo, il moto browniano può essere utilizzato per introdurre una componente stocastica nei movimenti degli agenti. Questa componente stocastica permette una migliore esplorazione dello spazio delle soluzioni, contribuendo a evitare il problema della convergenza a soluzioni subottimali. In particolare, l'introduzione di una variazione casuale nei movimenti degli agenti consente loro di esplorare aree più ampie dello spazio di ricerca, aumentando così la probabilità di trovare soluzioni globali ottimali. Ad esempio, l'algoritmo di ottimizzazione basato sul moto browniano dei gas (GBMO) combina i principi fisici del moto browniano con strategie di ottimizzazione per migliorare la ricerca di soluzioni ottimali in spazi complessi (7). Altri studi dimostrano che il moto browniano può essere combinato con altri algoritmi di ottimizzazione, come l'algoritmo del volo della farfalla e l'ottimizzazione degli stormi di particelle, per migliorare ulteriormente le prestazioni e la convergenza (8), (9).

La componente browniana $\eta_i(t)$ è definita come:

$$\eta_i(t) = \begin{bmatrix} \text{gauss}_i(t) \cdot \cos(\theta_i(t)) \\ \text{gauss}_i(t) \cdot \sin(\theta_i(t)) \end{bmatrix} \quad (5)$$

dove:

$$\theta_i(t) \sim \mathcal{U}(0, 2\pi) \quad (6)$$

rappresenta la direzione casuale, distribuita uniformemente nell'intervallo $[0, 2\pi]$.

$$\text{gauss}_i(t) \sim \mathcal{N}(\mu, \sigma^2) \quad (7)$$

rappresenta l'intensità del moto browniano, distribuita secondo una distribuzione gaussiana con media μ e varianza σ^2 .

Questa componente aggiuntiva introduce una variazione casuale nel movimento degli agenti, permettendo una migliore esplorazione dello spazio e contribuendo a evitare che gli agenti si raggruppino in posizioni subottimali.

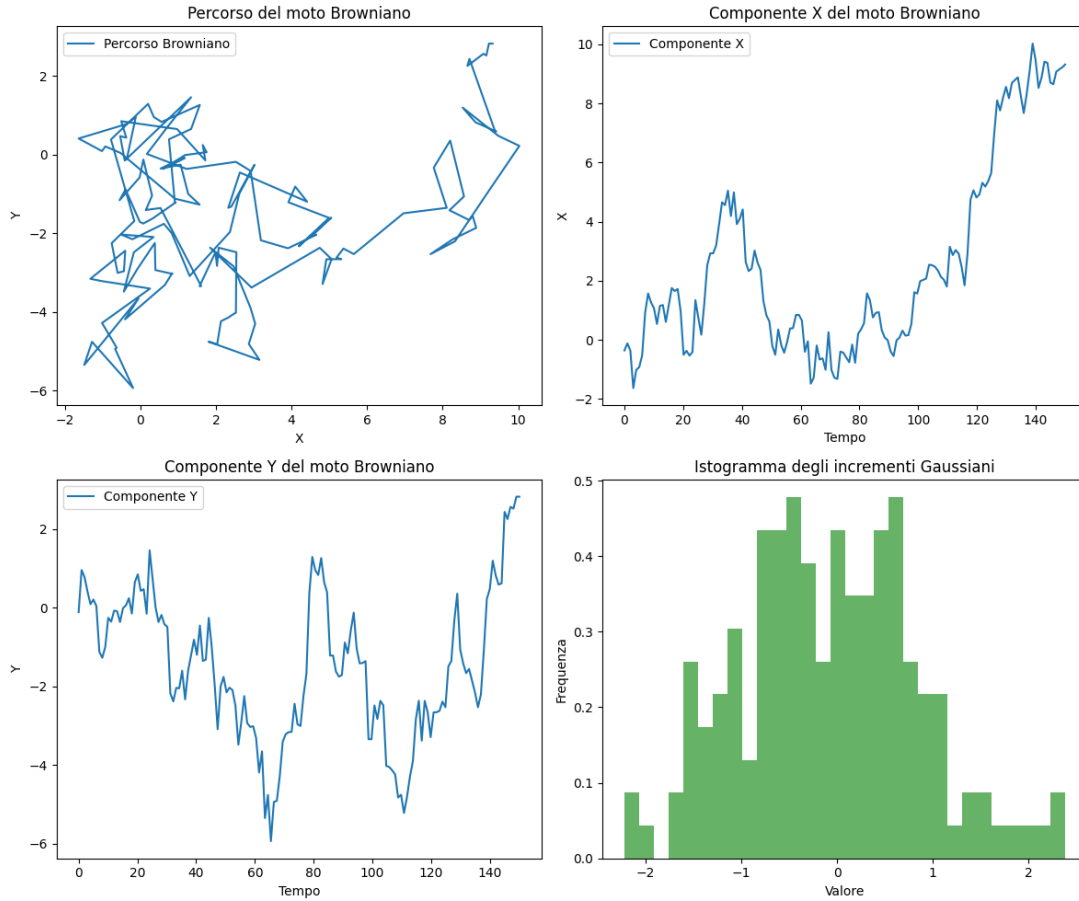


Figura 5: Grafici del moto Browniano

La Figura 5 illustra diverse caratteristiche del moto browniano. Il grafico in alto a sinistra mostra la traiettoria in due dimensioni di una particella che segue un moto browniano, evidenziando la natura casuale e imprevedibile del percorso. I grafici in alto a destra e in basso a sinistra mostrano rispettivamente le componenti X e Y del moto browniano nel tempo, che evidenziano come queste componenti variano casualmente. Infine, l'istogramma in basso a destra rappresenta la distribuzione degli incrementi gaussiani, confermando che questi incrementi seguono una distribuzione normale con media zero.

2.3 Algoritmo V3: aggiunta di un potenziale repulsivo

3 Implementazione degli algoritmi

3.1 Descrizione del codice (e tabelle con pseudocodice)

3.2 Librerie Python utilizzate

4 Simulazione e discussione dei risultati

5 Conclusioni

Riferimenti bibliografici

- [1] J. Cortes, S. Martinez, T. Karatas and F. Bullo, "Coverage control for mobile sensing networks," in IEEE Transactions on Robotics and Automation, vol. 20, no. 2, pp. 243-255, April 2004, doi: 10.1109/TRA.2004.824698. keywords: Sensor phenomena and characterization;Remotely operated vehicles;Mobile robots;Temperature sensors;Biosensors;Animals;Optimization methods;Partitioning algorithms;Prototypes;Infrared sensors, <https://ieeexplore.ieee.org/document/1284411>
- [2] S. Meng and Z. Kan, "Deep Reinforcement Learning-Based Effective Coverage Control With Connectivity Constraints," in IEEE Control Systems Letters, vol. 6, pp. 283-288, 2022, doi: 10.1109/LCSYS.2021.3070850. keywords: Sensors;Monitoring;Dynamics;Task analysis;Reinforcement learning;Vehicle dynamics;Training;Coverage control;deep reinforcement learning;multi-agent systems, <https://ieeexplore.ieee.org/document/9395182>
- [3] Hübel, Nico. (2008). Coverage Control with Information Decay in Dynamic Environments. IFAC Proceedings Volumes. 41. 4180-4185. 10.3182/20080706-5-KR-1001.00703. https://www.researchgate.net/publication/254995455_Coverage_Control_with_Information_Decay_in_Dynamic_Environments
- [4] Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep Learning. MIT Press. <https://www.deeplearningbook.org/>
- [5] Stipanović, D.M., Hokayem, P.F., Spong, M.W., & Šiljak, D.D. (2007). Cooperative avoidance control for multi-agent systems. *Journal of Dynamic Systems, Measurement, and Control*, 129, 699-707. doi:10.1115/1.2764502. <https://asmedigitalcollection.asme.org/dynamicsystems/article-abstract/129/5/699/446752/Cooperative-Avoidance-Control-for-Multiagent>
- [6] E. Nelson, *Dynamical Theories of Brownian Motion*, Princeton University Press, 1967. <https://web.math.princeton.edu/~nelson/books/bmotion.pdf>
- [7] M. Abdechiri, M. R. Meybodi, H. Bahrami, *Gases Brownian motion optimization: an algorithm for optimization (GBMO)*, Applied Soft Computing, 2013. <https://www.sciencedirect.com/science/article/pii/S1568494612001834>
- [8] K. Sowjanya, S. K. Injeti, *Investigation of butterfly optimization and gases Brownian motion optimization algorithms for optimal multilevel image thresholding*, Expert Systems with Applications, 2021. <https://www.sciencedirect.com/science/article/pii/S095741742100717X>
- [9] Q. Xiaohong, Q. Xiaohui, *An evolutionary particle swarm optimizer based on fractal Brownian motion*, Journal of Computational Intelligence, 2012. <https://www.ingentaconnect.com/contentone/asp/jcies/2012/00000001/00000001/art00023>