# GIỚI THIỆU

## Tổng quan đề tài

Đề tài trình bày những vấn đề liên quan đến việc thiết kế hệ thống định vị trong nhà. Định vị là việ c xác định vị trí của đối tượng trong một khu vực được toạ độ hoá bởi một hệ quy chiếu cho trước. Một hệ thống định vị hoạt động trên một thiết bị điện toán phải có chức năng xác định vị trí của thiết bị trong khu vực cho trước với một độ chính xác nhất định.

Một hệ thống định vị trong nhà là một hệ thống có thể xác định vị trí của đối tượng hoặc con người trong một toà nhà sử dụng các tín hiệu sóng radio, bức xạ từ trường, hoặc các dữ liệu cảm biến từ các thiết bị di động. Trên thực tế, đã có rất nhiều nghiên cứu để phát triển các hệ thống định vị trong nhà sử dụng các công nghệ khác nhau với các mục tiêu khác nhau như giảm chi phí, tăng độ chính xác của kết quả định vị. Nhờ vậy, nhiều nghiên cứu trong lĩnh vực này đã được triển khai, các phát minh nhằm cải thiện các hệ thống định vị trong nhà và tăng tính ứng dụng của chúng.

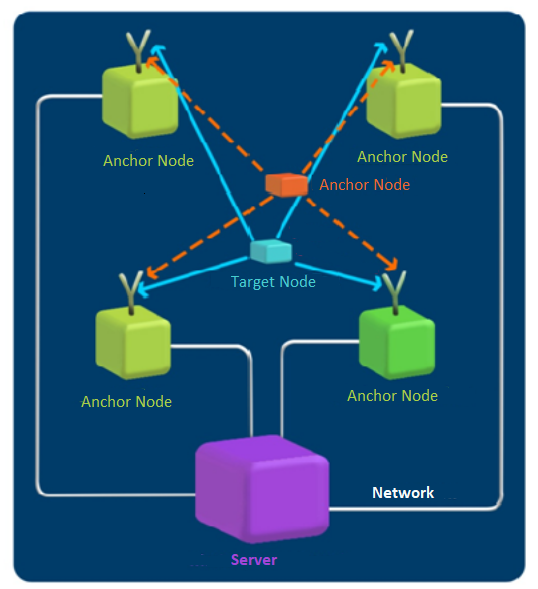
Hệ thống này đã và đang được sử dụng rộng rãi trong các nhà xưởng, công ty,… Điển hình là các ứng dụng theo dõi vật thể, robot, quản lý tài sản, chấm công,…

Cấu trúc cơ bản của một hệ thống định vị trong nhà bao gồm các thành phần sau:

**Anchor Node**: là các thành phần được đặt cố định trong không gian (nút neo). Các điểm này được xem như là các điểm tham chiếu của hệ thống, giúp chúng ta cố định được hệ tọa độ trong không gian.

**Target Node**: là các thành phần di chuyển trong không gian. Nhiệm vụ của đề tài này là tìm ra được vị trí của các Node này sao cho sai số nằm trong mức cho phép (3m - ổn, 1m – tốt), độ ổn định cao.

**Server**: là hệ thống lưu trữ các thông tin từ Anchor Node và Target Node đưa lên và sử dụng các thông tin đó để tính toán, hiển thị kết quả, điều khiển.

Trong luận văn này, Anchor Node – Target Node lần lượt là thiết bị Thu – Phát tín hiệu hoặc ngược lại.

Hình 1.1 Cấu trúc cơ bản của Indoor Positioning System

## Các thách thức của hệ thống định vị trong nhà

Khi xây dựng một hệ thống định vị, điều đầu tiên cần quan tâm là khả năng xác định vị trí đối tượng của hệ thống. Hệ thống cần xác định vị trí của đối tượng trong khoảng sai số cho phép, tuỳ trường hợp cụ thể.

Một yếu tố quan trọng khác trong đánh giá hiệu năng của hệ thống là độ trễ, hệ thống phải có khả năng định vị đối tượng trong một khoảng thời gian cho phép thì kết quả định vị của hệ thống mới có giá trị.

Mặt khác, chi phí triển khai hệ thống cũng là một yếu tố quan trọng cần cân nhắc. Các hệ thống định vị trong nhà hiện có thường yêu cầu cài đặt các thiết bị đắt tiền làm các mốc định vị, điều này góp phần làm tăng chi phí triển khai hệ thống và khó có thể áp dụng rộng rãi. Ngược lại, các hệ thống không yêu cầu các thiết bị cài đặt sẵn thường đưa ra các kết quả có độ chính xác không cao.

## Các kỹ thuật định vị trong nhà

Khi tín hiệu vệ tinh GPS không khả thi để định vị trong nhà, các hệ thống định vị trong nhà cần lựa chọn các tín hiệu khác phục vụ cho việc định vị. Sự phổ biến của các công nghệ không dây như Wi-Fi hay Bluetooth mở ra khả năng ứng dụng các tín hiệu này cho việc định vị trong nhà.

Bên cạnh đó, các thuật toán định vị cũng là một phần không thể thiếu trong các hệ thống định vị trong nhà. Phần tiếp theo của luận văn sẽ trình bày về các loại tín hiệu thường được sử dụng trong các hệ thống định vị trong nhà cũng như các thuật toán định vị phổ biến.

### Các công nghệ không dây

Các công nghệ không dây được sử dụng rất phổ biến trên toàn cầu nhằm mục đích liên lạc và chia sẻ dữ liệu. Các hệ thống định vị trong nhà hiện có thường dựa trên các công nghệ không dây được triển khai tại môi trường hoạt động.

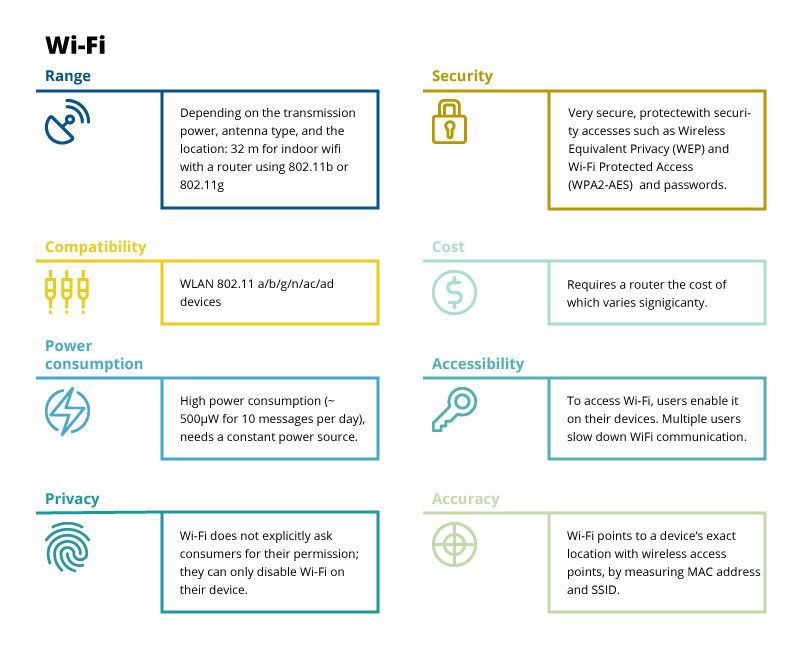
Tuy các công nghệ này chủ yếu được thế kế nhằm mục đích liên lạc và truy cập dữ liệu, không nhằm mục đích định vị, các nhà nghiên cứu vẫn có thể dựa trên các thuộc tính thu nhận được từ tín hiệu vô tuyến để ước lượng vị trí của nguồn phát tín hiệu từ đó ứng dụng vào các hệ thống định vị trong nhà.

#### Công nghệ WiFi

**Wi-Fi** là một công nghệ mạng không dây sử dụng sóng radio để cung cấp các kết nối Internet và mạng tốc độ cao dựa trên các tiêu chuẩn IEEE 802.11. Khi một dòng tần số vô tuyến được cung cấp cho ăng ten, nó sẽ tạo ra một trường điện từ có khả năng lan truyền trong lĩnh vực .

Hình 1.2 Đặc tính công nghệ WiFi

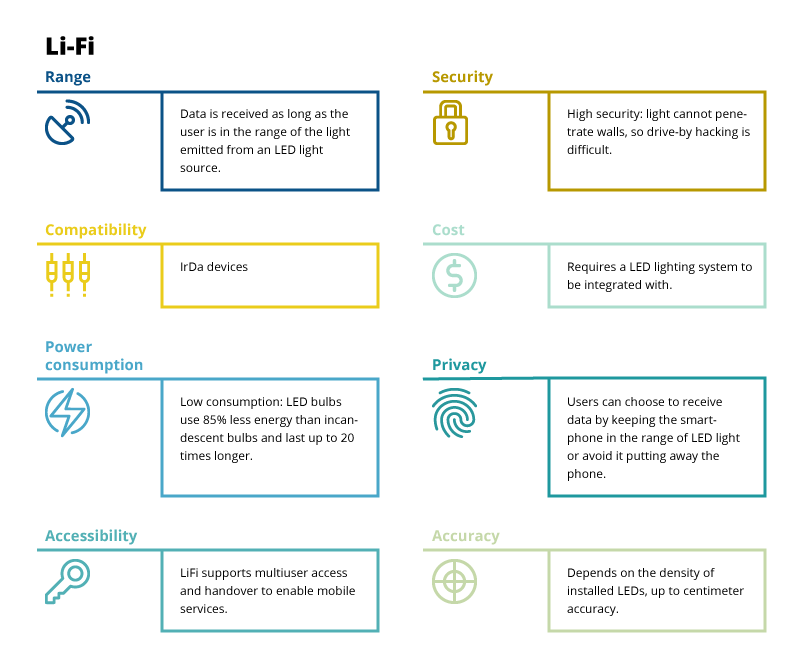
Các thiết bị được trang bị bộ điều hợp mạng không dây phát hiện tín hiệu không dây được phát bởi các điểm truy cập và điều chỉnh vào nó.



#### Công nghệ Light Fidelity (Li-Fi)

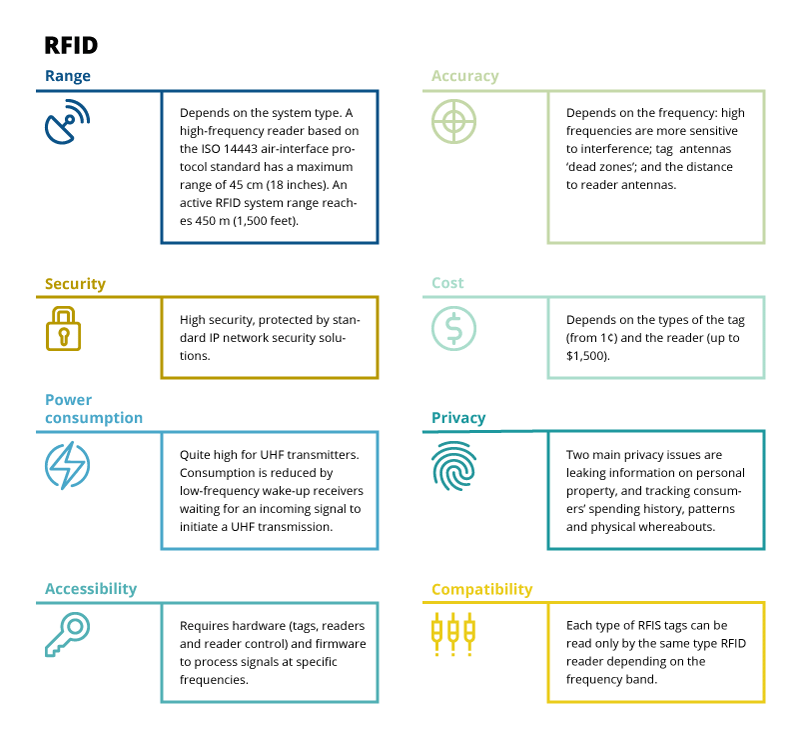
**Li-Fi** đề cập đến một hình thức không dây tương đối mới của công nghệ Giao tiếp ánh sáng có thể nhìn thấy, sử dụng các điốt phát sáng như một phương tiện để liên lạc tốc độ cao. Dữ liệu được truyền bằng cách điều chỉnh cường độ ánh sáng LED trong khoảng thời gian nano giây, quá nhanh để có thể phát hiện bằng mắt người.

Hình 1.3 Đặc tính công nghệ Li-Fi



#### Công nghệ Nhận dạng tần số vô tuyến (RFID)

**RFID** là công nghệ sử dụng sóng vô tuyến để đọc và thu thập thông tin được lưu trữ trên thẻ gắn với vật thể – một thiết bị điện tử nhỏ bao gồm một con chip nhỏ thường mang 2.000 byte dữ liệu trở xuống và ăng ten . Một thẻ có thể được đọc từ cách xa đến vài feet và không cần phải nằm trong tầm nhìn trực tiếp của người đọc để được theo dõi.

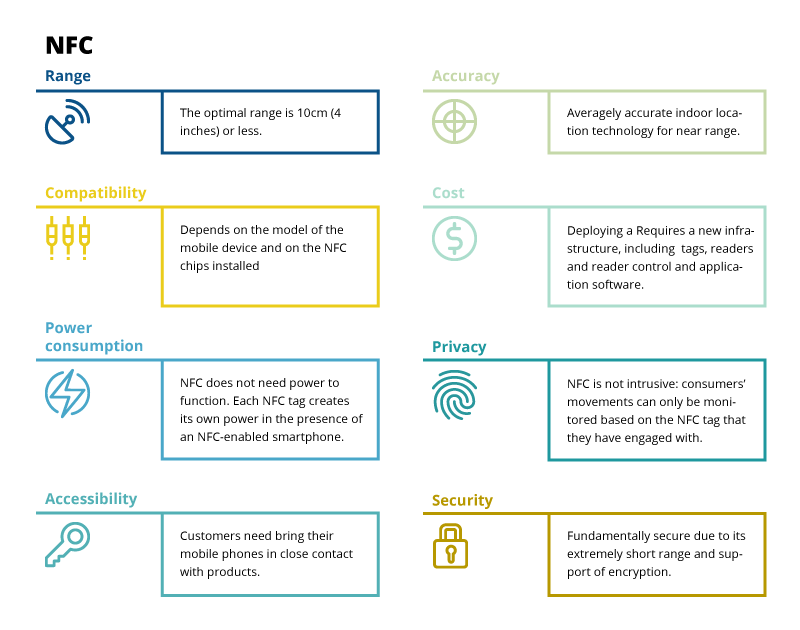


Hình 1.4 Đặc tính công nghệ RFID

#### Công nghệ Giao tiếp trường gần (NFC)

**NFC** là một liên kết không dây tầm ngắn, phát triển từ công nghệ RFID và có thể truyền một lượng nhỏ dữ liệu. Thẻ NFC chỉ giao tiếp với điện thoại thông minh có bật NFC khi chúng được đặt gần nhau (tối ưu dưới 4cm).

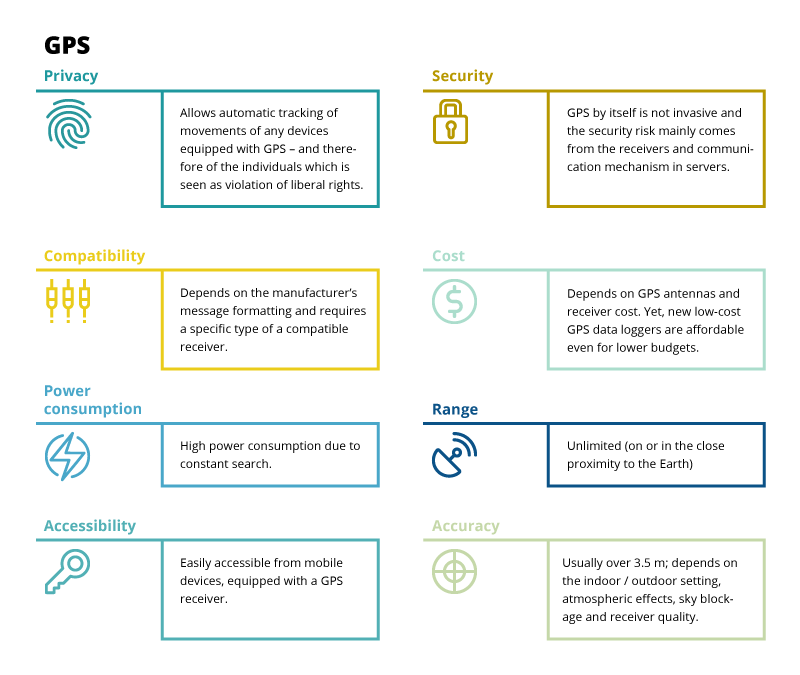
NFC có nhiều công dụng, đặc biệt là trong việc thanh toán, biến điện thoại thành ví điện tử. NFC có tiềm năng để thực hiện việc đó, loại bỏ thẻ tín dụng, séc (cheque) và các phương tiện thanh toán khác.



Hình 1.5 Đặc tính công nghệ NFC

#### Hệ thống Định vị toàn cầu (GPS)

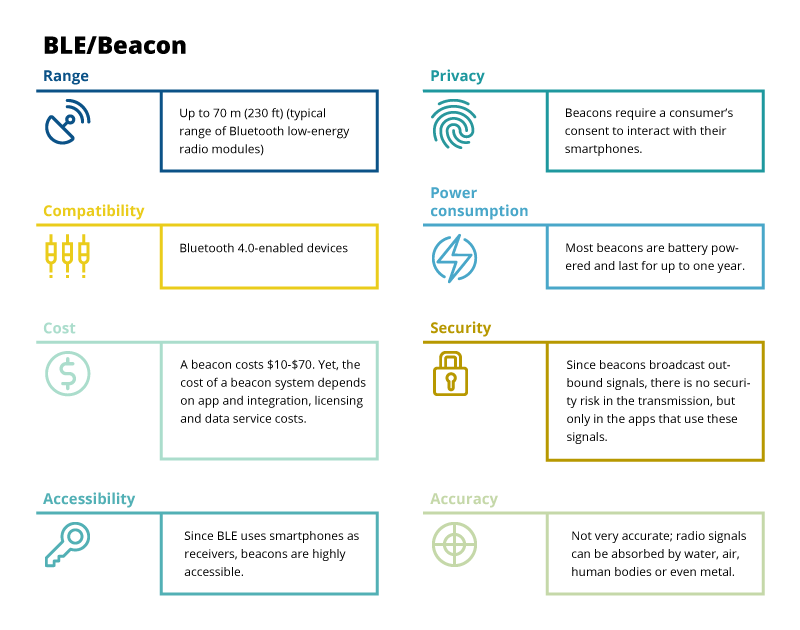
**GPS** là một hệ thống định vị dựa trên vệ tinh được tạo thành từ ít nhất 24 vệ tinh, thuộc sở hữu của chính phủ Hoa Kỳ và được điều hành bởi Không quân Hoa Kỳ. Nó cung cấp định vị địa lý và thông tin thời gian cho máy thu GPS ở bất cứ đâu trên hoặc gần Trái đất nơi có đường ngắm không bị cản trở đối với bốn hoặc nhiều vệ tinh GPS.



Hình 1.6 Đặc tính công nghệ GPS

#### Công nghệ Bluetooth Năng lượng thấp (BLE)

**BLE** là công nghệ mạng cá nhân không dây, được thiết kế để giảm mức tiêu thụ điện năng và chi phí của Bluetooth cổ điển trong ngành chăm sóc sức khỏe, thể dục, đèn hiệu, an ninh và giải trí gia đình. Beacon – một trong những thiết bị IoT phổ biến nhất dựa trên công nghệ BLE là một máy phát vô tuyến Bluetooth nhỏ; nó liên tục truyền một tín hiệu duy nhất mà các thiết bị khác có thể nhìn thấy. Đèn hiệu có thể phát tín hiệu vô tuyến được tạo thành từ sự kết hợp của các chữ cái và số được truyền trong khoảng thời gian đều đặn khoảng 1/10 giây.



Hình 1.7 Đặc tính công nghệ BLE

BLE phù hợp hơn để truyền một lượng nhỏ dữ liệu với tốc độ 1 Mb / giây, như đọc cảm biến về nhiệt độ, chi tiết gia tốc, tọa độ GPS, v.v.

So sánh với Chuẩn 802.11ac Wi-Fi có thể truyền với tốc độ lên tới 1,3 Gbps, vì vậy nó rất lý tưởng cho các tệp và dữ liệu lớn hơn. Wi-Fi Direct cung cấp tốc độ truyền dữ liệu tối đa, gấp khoảng 10 lần so với Bluetooth Classic. Nhưng BLE chậm hơn khoảng 2-3 lần so với Bluetooth Classic và chậm hơn 20-30 lần so với Wi-Fi Direct.

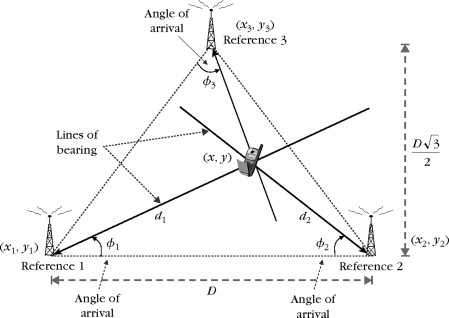
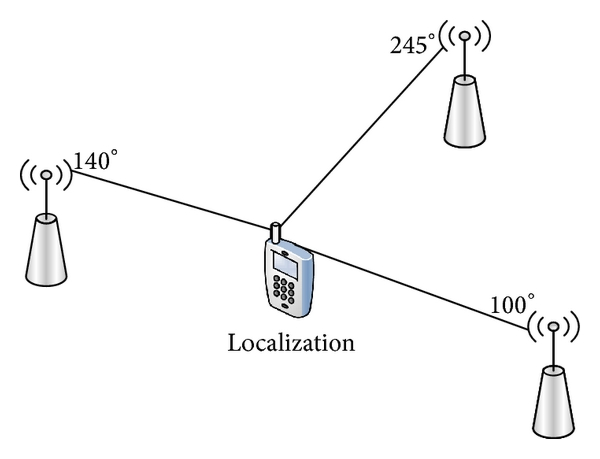
### Các Kỹ thuật Định vị

Song song cùng với các công nghệ không dây được sử dụng, ta cùng cần đến các phương pháp, thuật toán, kỹ thuật định vị phù hợp, từ đó chúng ta mới có thể hình thành được mô hình cơ bản của hệ thống, xác định được tiền đề cốt lõi để nghiên cứu và phát triển hệ thống một cách toàn diện. Các kỹ thuật định vị thường được sử dụng là **Angle of Arrival (AOA), Time Of Arrival (TOA), Time Diference of Arrival (TDOA), Received Signal Strength Indicator (RSSI),…**

#### Angle of Arrival (AOA)

Góc đến (Angle of Arrival – AOA) của một tín hiệu là hướng mà tín hiệu đó được nhận bởi máy thu. Việc đo lường AOA được thực hiện bằng cách xác định hướng lan truyền của sóng Radio trên 1 dải antenna hoặc được xác định từ cường độ tín hiệu tối đa đo được trong quá trình quay antenna.

Khi ta có được ít nhất AOA của 2 trạm thu nhận sóng Radio, vị trí của thiết bị cần tìm chính là giao điểm của 2 đường thẳng xuất phát từ 2 trạm thu nhận đó, có góc quay tương ứng với AOA của mỗi trạm.



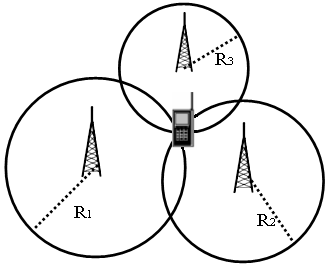
Hình 1.8 Mô hình hệ thống sử dụng kỹ thuật Angle of Arrial

#### Time of Arrival (TOA)

Thời gian đến (Time of Arrival – TOA / Time of Flight – TOF) của tín hiệu là thời gian mà tín hiệu bay trong không gian, nói cách khác là thời gian từ khi được phát ra cho đến khi được thu lại. Như ta đã biết, vận tốc của tín hiệu bay trong không gian chính bằng vận tố ánh sáng (299.792.458 m/s), vậy nếu ta có được tổng thời gian bay thì ta có thể xác định được quãng đường tín hiệu đã đi bằng cách lấy tổng thời gian bay chia cho vận tốc ánh sáng.

Với mỗi quãng đường mà ta tính được ở mỗi trạm thu, ta có thể vẽ đường 1 hình tròn bán kính bằng quãng đường đó là tập hợp những điểm có thể là vị trí của thiết bị, vậy nếu ta vẽ được 2 vòng tròn, tập hợp đso chỉ còn 2 điểm, khi ta có 3 vòng tròn ta sẽ có vị trí chính xác của thiết bị. Nếu ta xác định vị trí của thiết bị trong không gian 3D thì sẽ cần đến ít nhất 4 mặt cầu.

Kỹ thuật này yêu cầu các trạm thu phải được đồng bộ về mặt thời gian.



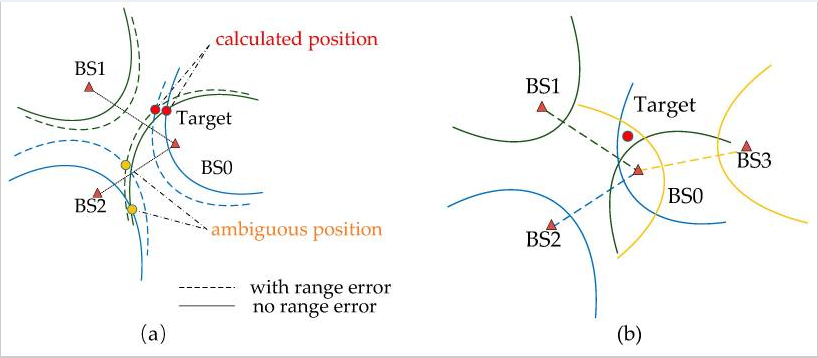
Hình 1.9 Mô hình hệ thống sử dụng kỹ thuật Time of Arrival

#### Time Difference of Arriaval (TDOA)

Sự khác biệt thời gian đến (Time Difference of Arrival – TDOA) là một phương pháp được phát triển từ TOA. Khi mà TOA yêu cầu đến sự đồng bộ về mặt thời gian giữa tất cả các trạm thu có sự chính xác đến nano giây, thì TDOA chỉ yêu cầu timer chính xác đến nano giây, sự đồng bộ về mặt thời gian giữa các trạm đã được loại bỏ, làm tối giản được hệ thống.

Mô hình TDOA giống như TOA, nhưng TDOA có thêm 1 trạm thu – phát cố định (Master Anchor), có trách nhiệm thu nhận các tín hiệu và phát lại ngay lập tức tín hiệu đó, giống như mô hình tiếng vang (ECHO).

Như vậy ngoài tín hiệu gốc, các trạm thu còn nhận thêm được tín hiệu ECHO có cùng nguồn gốc phát. Từ đây ta có thể xây dựng được hệ phương trình Hyperbolic hoặc Hyperboloids. Nghiệm của hệ này chính là vị trí cần tìm.

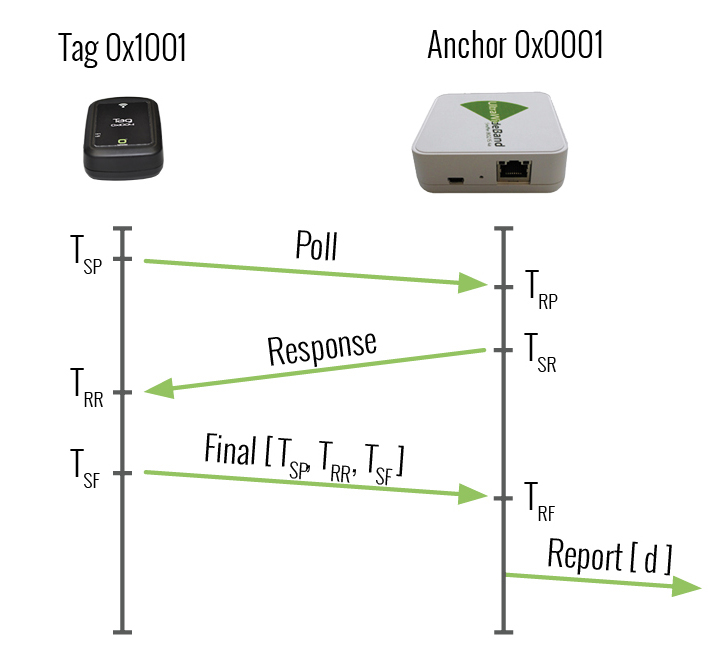


Hình 1.10 Mô hình hệ thống sử dụng kỹ thuật Time Difference of Arrival

#### Two-Ways Ranging (TWR)

Two-Ways Ranging cũng là 1 phương pháp được phát triển từ TOA. Mô hình TWR giống hệt như TOA, không sử dụng thêm 1 thiết bị ECHO như TDOA. Nhưng thay vì sử dụng One-Way Communication (Liên lạc một chiều) thì TWR sử dụng Two-Ways Communication (Liên lạc hai chiều), có nghĩa là ban đầu thiết bị bất định phát 1 đi 1 tín hiệu POLL, khi thiết bị cố định nhận được POLL, nó sẽ trả lại 1 tín hiệu RESPONSE, sau khi thiết bị bất định nhận được RESPONSE, nó sẽ phát đi 1 tín hiệu FINAL.

Đến đây ta đã có đủ thông tin cần để xác định khoảng cách của thiết bị bất định. Tập hợp nhiều khoảng cách tính được ta có được vị trí của thiết bị cần tìm. Phương pháp này cũng tối giản hệ thống khi không cần đồng bộ thời gian giữa các thiết bị.

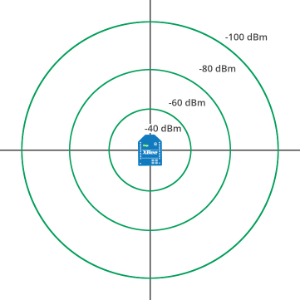


Hình 1.11 Mô hình hệ thống sử dụng kỹ thuật Time Difference of Arrival

#### Received Signal Strength Indicator (RSSI)

Chỉ số cường độ tín hiệu nhận được (Received Signal Strength Indicator RSSI) là một trong những phương pháp phổ biến nhất và đơn giản nhất để định vị. Lý do chính cho sự phổ biến của nó là việc tìm kiếm RSSI không yêu cầu phần cứng bổ sung và có thể được tìm thấy trên bất kỳ thiết bị nào sử dụng hầu hết mọi loại công nghệ truyền thông không dây. RSSI hoạt động bằng cách đo cường độ của các tín hiệu trên máy thu. Nó thường được sử dụng để tìm khoảng cách giữa máy phát và máy thu, bởi vì cường độ tín hiệu giảm đi trong môi trường khi tín hiệu truyền ra ngoài từ máy phát. Do tín hiệu lan truyền rất dễ bị nhiễu trong môi trường, RSSI thường dẫn đến các giá trị không chính xác có thể gây ra lỗi trong hệ thống định vị.

Mô hình hệ thống của RSSI tương tự như TDOA, chỉ cần tìm được các khoảng cách từ máy phát đến máy thu ta có thể tìm được vị trí chính xác của thiết bị.



Hình 1.12 Mô hình mô phỏng cường độ tín hiệu phụ thuộc vào khoảng cách

### Lựa chọn kỹ thuật định vị

Trong luận văn này sẽ sử dụng công nghệ Bluetooth Năng lượng thấp và kỹ thuật RSSI để xác định vị trí.

Ưu điểm:

* Ít tiêu tốn năng lượng, thích hợp cho việc tích hợp các ứng dụng IOT.
* Chi phí mua thiết bị, lắp đặt thấp. Tối ưu chi phí cho hệ thống. (Mọi smartphone hiện nay đều được tích hợp Bluetooth, thiết bị thu không đắt đỏ như router wifi)
* Không yêu cầu thêm phần cứng để đọc cường độ tín hiệu.
* Cách thức liên lạc, thuật toán đơn giản, không làm phức tạp hệ thống.

Nhược điểm:

* Tầm hoạt động thấp (~50m cho điều kiện lý tưởng, khoảng cách tối ưu ~20m), nếu xây dựng trong không gian rộng lớn cần rất nhiều thiết bị tạo thành mạng lưới, không ổn định khi chuyển vùng.
* Tín hiệu không ổn định, chịu tác động của môi trường khá lớn, tần số phát tin không cao (tối đa 10 tin / 1 giây) => cần các bộ lọc => tốc độ hội tụ thấp.



# Cấu trúc hệ thống

## Mô hình phần cứng

Theo Hình 1.1: Cấu trúc cơ bản của Indoor Positioning System, ta cần các thiết bị có nhiệm vụ thu, phát sóng Bluetooth Low Energy, ngoài ra thiết bị thu sóng phải có thêm tính năng kết nối mạng (WiFi hoặc dây Ethernet – để tăng tính linh động cho thiết bị ta sẽ sử dụng WiFi) để kết nối, gửi dữ liệu lên server. Như vậy, ta sẽ chọn các thiết bị sau:

### Kit RF Thu Phát Wifi BLE ESP32 NodeMCU LuaNode32

ESP32 là chip kết hợp WiFi và Bluetooth băng tần 2.4GHz được sản xuất trên tiến trình ultra-low-power 40nm của TSMC, được phát triển bởi ESPRESSIF SYSTEMS (SHANGHAI) CO., LTD.. CPU của ESP32 là Xtensa® single-/dual-core 32-bit LX6 microprocessor, 448 KB ROM, 520KB SRAM, xung nhịp cơ bản 8MHz, thạch anh ngoài 2~60MHz (40MHz cho chức năng WiFi/Bluetooth).

Kit RF thu phát Wifi BLE ESP32 NodeMCU LuaNode32 được phát triển trên nền module trung tâm là ESP32 với công nghệ Wifi, BLE và nhân ARM SoC tích hợp mới nhất hiện nay, với ưu điểm là cách sử dụng dễ dàng, ra chân đầy đủ, tích hợp mạch nạp và giao tiếp UART CP2102, Kit Wifi BLE ESP32 NodeMCU LuaNode32 là sự lựa chọn hàng đầu trong các nghiên cứu, ứng dụng về Wifi, BLE, IoT và điều khiển, thu thập dữ liệu qua mạng.

Tuy nhiên, ESP32 NodeMCU LuaNode32 sử dụng Antenna Onboard, đây là loại antenna có hướng, loại antenna này nếu thu nhận sóng khác hướng đến thì sẽ làm cường độ tín hiệu bị giảm đi đáng kể dẫn đến khả năng ước lượng khoảng cách thiết bị.

Vì vậy thay vì sử dụng Antenna Onboard, ta sẽ sử dụng Antenna Ipex là loại antenna lưỡng cực được cắm ngoài để hạn chế tối thiểu sự chênh lệch giữa các hướng đến của tín hiệu. Vậy ta sẽ sử dụng module ESP32 ESP-WROOM-32U.



Hình 2.1 Kit ESP32 NodeMCU LuaNode32 (trái)

ESP32 ESP-WROOM-32U (phải)

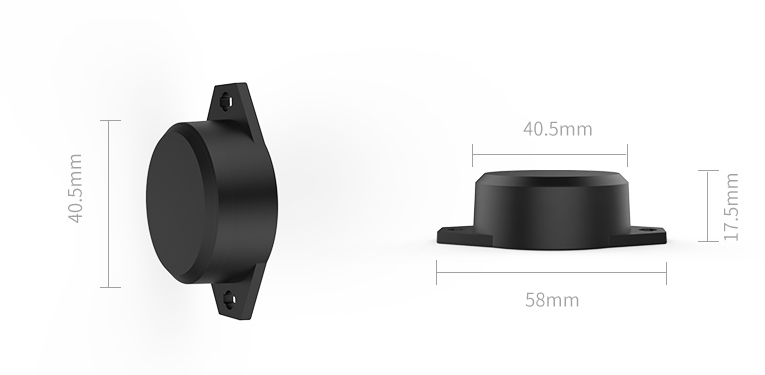
Khi sử dụng ESP32 ESP-WROOM-32U, để tiết kiệm công sức đặt Kit phát triển (hiện tại Việt Nam có/ít kinh doanh) hay thiết kế mạch phần cứng tích hợp, ta có thể sử dụng Kit ESP32 NodeMCU LuaNode32 (khá phổ biến tại Việt Nam) và thay module hiện có của kit bằng module ESP32 ESP-WROOM-32U.

### Beacon

Beacon là các thiết bị phát các tin Bluetooth Low Energy liên tục để ESP32 có thể read được RSSI của các tin này, từ đó gửi lên server để tính toán và ước lượng. Một chu kì phát tin của Beacon sẽ giao động từ 100-5000ms. Khi chu kỳ gửi càng ngắn (tức tần suất phát tin càng tăng) thì lượng dữ liệu được đưa lên server càng nhiều, dẫn đến đáp ứng về ướng lượng càng nhanh và chính xác hơn. Từ đó ta có thể xây dựng một hệ thống với thời gian thực (real time).

Hiện nay trên thị trường Việt Nam có rất nhiều loại Beacon đa dạng về mẫu mã chất lượng, giá thành,….

E9 Dear Beacon là một kiểu beacon sử dụng Bluetooth Low Energy, dựa trên chip BLE thế hệ mới – Nordic nRF52. Beacon này hỗ trợ cả công nghệ iBeacon(BLE của Apple) và Eddystone(BLE của Google), có thể tùy chỉnh cấu hình thông qua App trên Smart Phone.





Hình 2.2 Cấu trúc của E9 Dear Beacon

Dear Beacon được thiết kế cho cho các ứng dụng trong môi trường ẩm ướt, bởi vì nó đạt tiêu chuẩn chống nước IP67. Ngoài nó, Dear Beacon còn được trang bị thêm cảm biến gia tốc (accelerometer sensor) nhằm thực hiện nhiều loại ứng dụng như, theo dõi tài sản, quảng cáo tại danh lam thắng cảnh,….

Bên cạnh đó, có thể lắp đặt Dear Beacon với ốc vít tại các khu vực công cộng và dễ dàng được tháo lắp bới thiết kế xoay vỏ.

E9 Dear Beacon được trang bị viên pin 1000mAh có thể hoạt động liên tục 3 năm trong điều kiện Công suất phát 0dBm và chu kỳ 900ms). Tầm hoạt động có thể đạt đến 100 mét trong điều kiện lý tưởng và có thể tùy chỉnh công suất phát.

## Server Trung tâm

Các thiết bị phần cứng chỉ đóng vai trò thu thập dữ liệu để gửi lên server, mọi thao tác tính toán, lưu trữ đều được thực hiện một cách tự động trên server. Vì vậy ta phải chọn cấu hình của Server phù hợp với số lượng công việc mà Server phải giải quyết, cũng như lựa chọn phương thức liên lạc giữa server và thiết bị một cách phù hợp.

### Microsoft Azure Compute Engine

Với sự phát triển với tốc độ vũ bão của khoa học công nghệ như hiện nay thì việc ứng dụng Cloud Computing trở nên phổ biến và được nhiều doanh nghiệp cũng như các ứng dụng khác trong đời sống xã hội sử dụng là điều cần thiết.

Ưu điểm nổi bật nhất mà Cloud Computing mang đến là việc sử dụng và khai thác các tài nguyên trên máy tính hiệu quả hơn rất nhiều, điều này giúp tiết kiệm chi phí đầu tư ban đầu và giúp cho người dùng không còn phải quá lệ thuộc vào các cơ sở hạ tầng. Không quá khi cho rằng Cloud Computing đang dần trở thành xu hướng của thời đại.

Với xu hướng mới của thời đại, các ông lớn trong ngành công nghệ luôn tìm cách xây dựng và phát triển một cách độc đáo và mãnh mẽ chẳng hạn như Google có Google Cloud Platform, IBM có IBM Cloud, và đương nhiên Microsoft cùng với Microsoft Azure không phải là ngoại lệ.

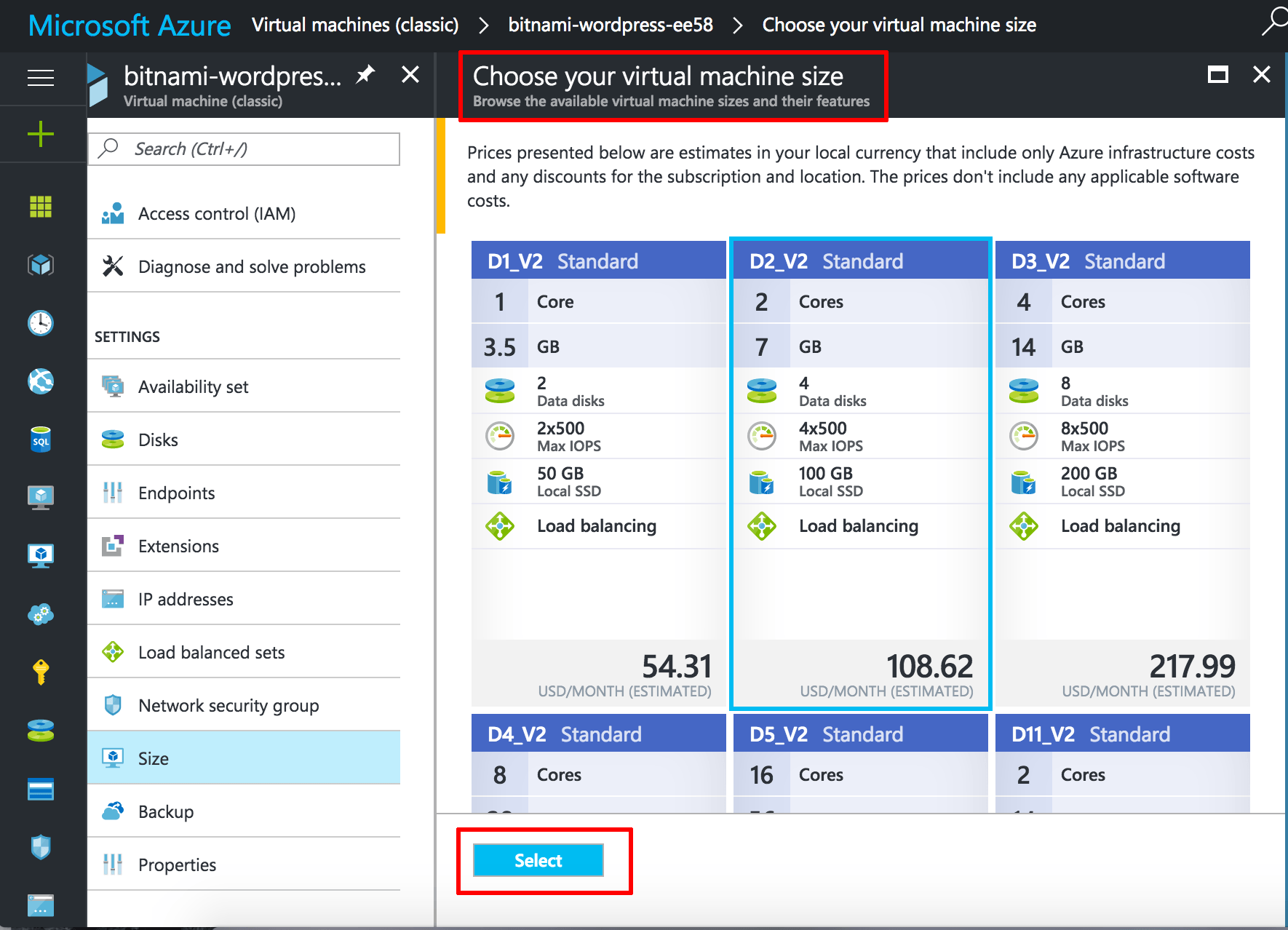
**Microsoft Azure** là một nền tảng ứng dụng cho công nghệ điện toán đám mây. Ta có thể sử dụng nền tảng này theo nhiều cách khác nhau. Chẳng hạn, ta có thể sử dụng Microsoft Azure để xây dựng các ứng dụng web để chạy hoặc lưu trữ dữ liệu trong Microsoft Azure DataCenters. Ngoài ra, chúng ta có thể dùng Microsoft Azure để tạo các máy ảo cho phát triển và kiểm thử phần mềm hoặc chạy SharePoints hay các nền tảng ứng dụng khác.

**Vitual Private Server** (VPS) của Microsoft Azure là một nền tảng mạnh mẽ, hỗ trơ nhiều tính năng, đáp ứng đủ điều kiện để chúng ta có thể thiết lập một server Indoor Positioning System. Một trong những tính năng đặc biệt của VPS là cho phép ta lựa chọn các thông số cấu hình phù hợp với yêu cầu của mình. Ta sẽ xây dựng một **Vitual Machine** (VM) làm VPS với cấu hình Intel Xeon E3-2 Core, Ram 4GB, IOPS 4000, chạy Hệ điều hành (OS) Ubuntu Server 18.4 LTS. VM này sẽ có nhiệm vụ thực hiện các phép tính toán trên dữ liệu mà các thiết bị phần cứng đưa lên, lưu trữ dữ liệu lên Database, thực hiện các Truy vấn (Command) của người dùng User.

Các ngôn ngữ sử dụng Python cho tính toán, JavaScript (Node.js) để thực hiện các truy vấn của User.



Hình 2.3 Microsoft Azure Virtual Machines



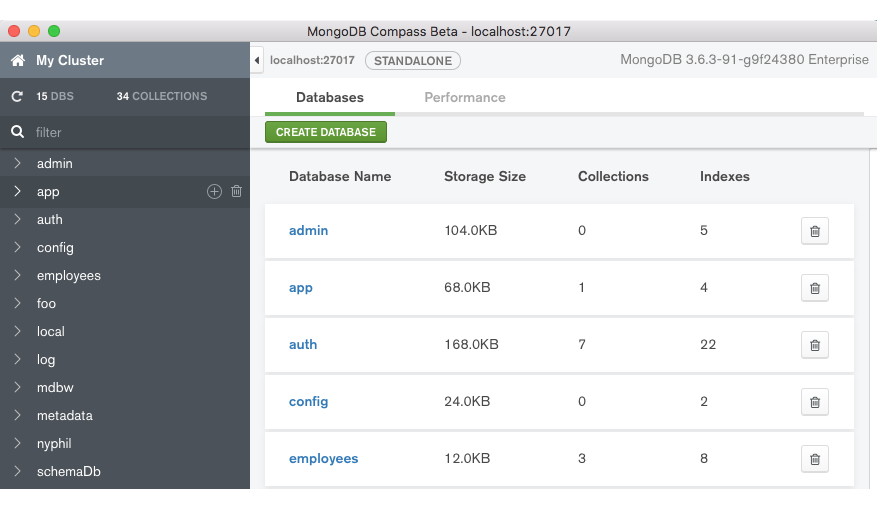
Hình 2.4 Tùy chọn cấu hình cho VM

### Cơ sở dữ liệu

**Cơ sở dữ liệu** (Database) là một tập hợp các dữ liệu có tổ chức, thường được lưu trữ và truy cập điện tử từ hệ thống máy tính. Khi cơ sở dữ liệu phức tạp hơn, chúng thường được phát triển bằng cách sử dụng các kỹ thuật thiết kế và mô hình hóa chính thức.

**Hệ quản lý cơ sở dữ liệu** (Database Management System - DBMS) là phần mềm tương tác với người dùng cuối, ứng dụng và chính cơ sở dữ liệu để thu thập và phân tích dữ liệu. Phần mềm DBMS bao gồm các tiện ích cốt lõi được cung cấp để quản trị cơ sở dữ liệu. Tổng cộng của cơ sở dữ liệu, DBMS và các ứng dụng liên quan có thể được gọi là "hệ thống cơ sở dữ liệu". Thông thường thuật ngữ "cơ sở dữ liệu" cũng được sử dụng để nói đến bất kỳ DBMS, hệ thống cơ sở dữ liệu hoặc ứng dụng nào được liên kết với cơ sở dữ liệu.

**MongoDB** là một cơ sở dữ liệu hướng tài liệu mã nguồn mở phổ biến được phát triển bởi 10gen, sau này được gọi là MongoDB Inc. Trong trường hợp này, các tài liệu được tạo và lưu trữ trong các tệp BSON, định dạng Binary JSON (JavaScript Object Notation),vì vậy tất cả các loại dữ liệu JS là được hỗ trợ. Trong trường hợp đó, MongoDB thường được áp dụng cho các dự án Node.js. Ngoài ra, JSON cho phép chuyển dữ liệu giữa các máy chủ và ứng dụng web bằng cách sử dụng định dạng có thể đọc được. Nó cũng là một lựa chọn tốt hơn, khi nói đến dung lượng và tốc độ lưu trữ, vì nó mang lại hiệu quả và độ tin cậy cao hơn.



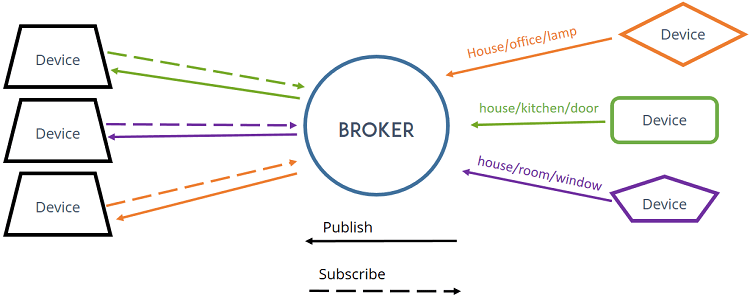
Hình 2.5 MongoDB

## Giao thức truyền thông MQTT

**Message Queuing Telemetry Transport** (MQTT) là một giao thức truyền thông điệp (message) theo mô hình publish/subscribe (xuất bản – theo dõi), sử dụng băng thông thấp, độ tin cậy cao và có khả năng hoạt động trong điều kiện đường truyền không ổn định.

MQTT là giao thức gọn nhẹ được thiết kế chủ yếu để kết nối các thiết bị bị hạn chế nguồn trên các mạng băng thông thấp. Mặc dù nó đã tồn tại trong hơn một thập kỷ nhưng chỉ khi có sự ra đời của Machine to Machine (M2M) và Internet of Things (IoT) mới làm cho nó trở thành một giao thức phổ biến.

MQTT sử dụng mẫu pub/sub để kết nối các bên với nhau. MQTT làm điều đó bằng cách tách người gửi (nhà xuất bản) với người nhận (người đăng ký). Nhà xuất bản gửi tin nhắn đến một chủ đề trung tâm, cái mà có nhiều người đăng ký đang chờ nhận tin nhắn. Nhà xuất bản và người đăng ký tự chủ, có nghĩa là họ không cần biết sự hiện diện của nhau.



Hình 2.6 Tổng quan giao thức MQTT

Các thuật ngữ trong MQTT:

**Client**: Bất kỳ nhà xuất bản hoặc người đăng ký nào kết nối với nhà môi giới tập trung qua mạng đều được coi là khách hàng. Cả nhà xuất bản và người đăng ký đều được gọi là khách hàng vì họ kết nối với dịch vụ tập trung, khách hàng có thể liên tục hoặc tạm thời. Khách hàng liên tục duy trì một phiên làm việc với nhà môi giới trong khi khách hàng tạm thời không được nhà môi giới theo dõi. Khách hàng thường kết nối với nhà môi giới thông qua thư viện và SDK. Có rất nhiều thư viện có sẵn cho C, C ++, Go, Java, C #, PHP, Python, Node.js và Arduino.

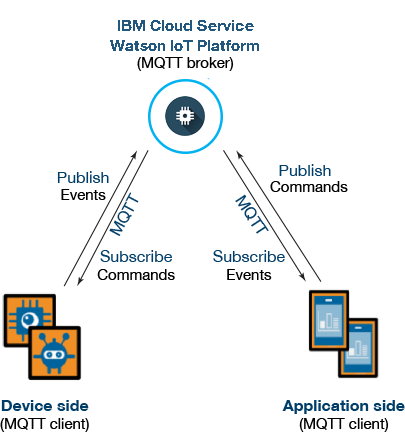
**Broker**: Người môi giới là phần mềm nhận tất cả các tin nhắn từ các khách hàng xuất bản và gửi chúng đến các khách hàng đăng ký. Nó giữ kết nối với các khách hàng liên tục. Tùy thuộc vào người triển khai để quyết định cách tạo lớp môi giới có thể mở rộng. Một số triển khai thương mại của các nhà môi giới MQTT bao gồm HiveMQ, Xively, AWS IoT và Loop.

**Topic**: Một chủ đề trong MQTT là điểm cuối mà khách hàng kết nối. Nó hoạt động như nơi phân phối trung tâm để xuất bản và đăng ký tin nhắn. Trong MQTT, một chủ đề là một vị trí nổi tiếng cho nhà xuất bản và người đăng ký. Nó được tạo ra khi chúng ta thiết lập kết nối với nhà môi giới. Chủ đề là các chuỗi phân cấp đơn giản, được mã hóa bằng UTF-8, được phân cách bằng dấu gạch chéo. Người đăng ký có thể chọn đăng ký một chủ đề cụ thể hoặc tất cả các chủ đề phụ thông qua các ký tự đại diện.

**Connection**: MQTT có thể được sử dụng bởi các máy khách dựa trên TCP/IP. Cổng tiêu chuẩn được giới thiệu bởi các công ty môi giới năm 1883, không phải là một cổng an toàn. Những nhà môi giới hỗ trợ TLS/SSL thường sử dụng cổng 8883. Để liên lạc an toàn, khách hàng và nhà môi giới dựa vào chứng chỉ kỹ thuật số.

Hiện nay có rất nhiều nhà cung cấp cho phép chúng ta xây dựng từ những server MQTT mức độ đơn giản cho đến phức tạp. Tùy vào cấu trúc của hệ thống và số lượng các kết nối, tần suất tin nhắn mà nhà cung cấp đề xuất cấu hình và mức giá khác nhau. Các nhà cung cấp điển hình như Microsoft Azure, Google Cloud, IBM Watson,….

Trong đề tài luận văn này chúng ta sẽ sử dụng MQTT server của IBM Watson vì ưu điểm là hệ thống băng thông lớn, giới hạn số lượng tin nhắn trong một ngày lớn (400.000 tin/ ngày so với 8000 tin/ngày của Microsoft Azure - ở gói Free). Giả sử ta có 1 thiết bị, 1 giây gửi 1 tin thì tổng cộng 1 giờ sẽ có 3600 tin và 1 ngày có 86400 tin, lớn hơn giới hạn 8000 tin/ngày của Microsoft Azure.



Hình 2.7 IBM Watson MQTT

# Phương trình Cường độ tín hiệu phụ thuộc Khoảng cách

## RSSI Distance Model

Các mô hình suy hao phổ biến bao gồm các mô hình suy hao trong không gian tự do, mô hình suy hao theo khoảng cách logarith. Các nghiên cứu đã chỉ ra rằng đặc tính suy hao của tín hiệu tuân theo phân phối xác suất Loga chuẩn (Lognormal Distribution). Đo khoảng cách dựa trên RSSI thường sủ dụng mô hình suy hao theo khoảng các logarith. Phương trình cơ bản của hệ thống như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (3.1) |

Trong đó:

là khoảng cách giữa Beacon và ESP ở khoảng cách .

là tham số suy hao.

là khoảng cách tham chiếu.

ξ là sai số ngẫu nhiên phân phối Gaussian, trung bình 0, phương sai .

Để thuận tiện trong tính toán, cho khi đo (1) trở thành:

|  |  |  |
| --- | --- | --- |
|  |  | (3.2) |

**Tham số suy hao** (PathLoss Exponent) là tham số trong sự suy hao về năng lượng của sóng điện từ khi nó lan truyền trong không gian. Pathloss Exponent 1 phần phản ánh được sự tác động của môi trường đến khả năng lan truyền của tín hiệu. Với mỗi môi trường khác nhau thì Pathloss Exponent sẽ khác nhau.

|  |  |  |  |
| --- | --- | --- | --- |
| Environment | Exponent | Environment | Exponent |
| Free space | 2 | In Building line of site | 1.6-1.8 |
| Urban area cellualr radio | 2.7-3.5 | Obstructed in building | 4-6 |
| Shadowed urban cellualr radio | 3-5 | Obstructed in Factories | 2-3 |

Bảng 3.1 Tham số Pathloss tương ứng với từng môi trường

## RSSI Position Model

Với phương trình (3.2) ta có được khoảng cách từ thiết bị đến điểm tham chiếu. Vậy nếu có nhiều điểm tham chiếu cho thiết bị đó ta sẽ dễ dàng xác định được vị trí của thiết bị so với các điểm tham chiếu. Nếu ta xác định trên không gian 2D thì sẽ cần ít nhất 3 phương trình, còn nếu là không gian 3D sẽ là 4 phương trình.

Giả sử ta có n điểm tham chiếu với vị trí đã bết trong không gian, nếu ta biết được khoảng cách tương ứng của thiết bị đến n điểm tham chiếu đó thì ta có thể thành lập hệ phương trình như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (3.3) |

Trong đó

là vị trí của thiết bị.

là vị trí tham chiếu thứ.

là số vị trí tham chiếu.

là khoảng cách tham chiếu thứ .

Hệ phương trình (3.3) là hệ phương trình đường tròn, để giải hệ phương trình này, ta chọn 1 phương trình làm gốc, sau đó lấy tất cả phương trình còn lại trừ cho phương trình đó. Giả sử ta chọn phương trình thứ n, khi đó ta có:

|  |  |  |
| --- | --- | --- |
|  |  | (3.4) |

Hệ trên có phương trình. Mỗi phương trình trong hệ (3.4) đều có dạng:

|  |  |  |
| --- | --- | --- |
|  |  | (3.5) |

Với .

Triển khai (3.5) ta được:

|  |  |  |
| --- | --- | --- |
|  |  | (3.6) |

Viết lại phương trình (3.6) theo dạng ta được:

|  |  |  |
| --- | --- | --- |
|  |  | (3.7) |

Vậy nếu viết hệ phương trình (3,4) theo dạng thì ta sẽ có như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (3.8) |
|  |  | (3.9) |

Sau khi biến đổi từ hệ phương trình (3.4) thành phương trình (3.8) ta thấy các phương trình trong hệ giờ đã trở thành phương trình bậc nhất hai biến số. Giải hệ phương trình này dễ hơn giải phương trình (3.4).

Kết quả của hệ phương trình (3.8) có thể được tính toán bằng công thức sau:

|  |  |  |
| --- | --- | --- |
|  |  | (3.10) |

Bởi vì luôn có tác động nhiễu từ các phép đo cũng như môi trường nên chỉ là kết quả gần đúng của hệ (3.4). Vì vậy ta gọi là độ lệch khoảng cách của từng phương trình. Ta xem như là 1 thước đo để xác định độ chính xác của kết quả bài toán.

# Áp dụng các giải thuật tối thiểu sai số

Mục đích của chúng ta là tìm vị trí gần với vị trí nhất để có thể theo dõi, quản lý, vì vậy thực hiện phép toán phương trình (3.10) là chưa đủ bởi vì hệ phương trình (3.8) chỉ là phương trình suy ra từ hệ phương trình (3.4), có nghĩa là đáp số của hệ (3.8) chưa chắc là kết quả tối ưu nhất của hệ (3.4) vì vậy ta phải sử dụng thêm các giải thuật khác để tìm được kết quả tối ưu.

Trong chương này ta sẽ tìm hiểu và áp dụng các thuật toán để tối ưu sai số như bộ lọc **Kalman**, phương pháp **Pathloss Exponent Improvement** và giải thuật **Particle Swarm Optimization**.

## Kalman Filter

Bộ lọc Kalman, được Rudolf (Rudy) E. Kálmán công bố năm 1960, là thuật toán sử dụng chuỗi các giá trị đo lường, bị ảnh hưởng bởi nhiễu hoặc sai số, để ước đoán biến số nhằm tăng độ chính xác so với việc sử dụng duy nhất một giá trị đo lường. Bộ lọc Kalman thực hiện phương pháp truy hồi đối với chuỗi các giá trị đầu vào bị nhiễu, nhằm tối ưu hóa giá trị ước đoán trạng thái của hệ thống.

Bộ lọc Kalman được ứng dụng rộng rãi trong kỹ thuật, phổ biến trong các ứng dụng định hướng, định vị và điều khiển các phương tiện di chuyển. Ngoài ra, bộ lọc Kalman còn được ứng dụng để phân tích dữ liệu trong các lĩnh vực xử lý tín hiệu và kinh tế.

Ứng dụng đầu tiên và nổi tiếng nhất chính là bộ lọc Kalman đã được áp dụng để điều hướng cho Dự án Apollo, trong đó yêu cầu ước tính quỹ đạo của tàu vũ trụ có người lái lên Mặt trăng và quay trở lại Trái đất.

Mặc dù Bộ lọc Kalman được ứng dụng trong nhiều lĩnh vực, nhưng nó được sử dụng chủ yếu với 2 mục đích chính:

* **Estimating the state of dynamic system** (Ước tính trạng thái của hệ thống động) – trong đó, hệ thống động là hệ thống có trạng thái thay đổi theo thời gian, mà trong vũ trụ này thì hiếm có thứ nào hoàn toàn bất biến. Từ những thông tin chứa đầy nhiễu và sự không chắc chắn (noise & uncertainty), bộ lọc Kalman có thể cung cấp cho chúng ta các giá trị ước tính (chính xác nhất có thể) về trạng thái hiện tại của hệ thống.
* The Analysis of Estimation Systems (Phân tích hệ thống dự đoán) Với các thông tin, giá trị ước tính về trạng thái hiện tại của hệ thống, bộ lọc kalman còn có thể tiên đoán được các giá trị ở trạng thái tiếp theo của hệ thống.

Trong mô hình IPS này, chúng ta sẽ sử dụng Kalman Filter vào dữ liệu RSSI đo được để hạn chế sai số và nhiễu ảnh hưởng. Ta có mô hình cơ bản như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (4.1) |

Trong đó

là vector trạng thái (trong trường hợp này là RSSI) tại điểm .

là ma trận mô hình chuyển đổi trạng thái.

là ma trận mô hình điều khiển đầu vào.

là vector điều khiển tại thời điểm .

vector ngẫu nhiên nhiễu hệ thống tại thời điểm .

Cùng với vector đo lường, ta có đủ phương trình cần thiết cho Kalman Filter:

|  |  |  |
| --- | --- | --- |
|  |  | (4.2) |

Trong đó

là vector đo lường (RSSI đo được) tại thời điểm .

là ma trận mô hình quan sát.

là vector ngẫu nhiên nhiễu đo lường.

Vì chúng ta đang sử dụng hệ thống tĩnh, không có sự tác động vào đối tượng nên vector điều khiển sẽ bằng 0 (). Ở đây, chúng ta áp dụng Kalman Filter vào RSSI thô đo được, nên vector trạng thái và vector đo lường sẽ chỉ có 1 phần tử, và sẽ bằng 1. Vậy ta có thể viết lại phương trình (4.1) và (4.2) như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (4.3) |
|  |  | (4.4) |

Ngoài ra, ta có hàm phân phối nhiễu hệ thống và đo lường như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (4.5) |

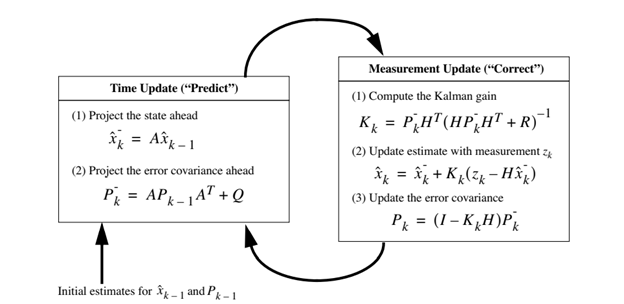
Trong đó

là hàm phân phối Gaussion.

là hiệp phương sai (covariance) nhiễu hệ thống.

là hiệp phương sai nhiễu đo lường.

Ở đây, chúng ta không đi sâu vào cách hình thành từng bước Kalman Filter, nên ta sẽ đi luôn vào cách sử dụng Kalman Filter như hình dưới.



Hình 4.1 Mô hình hoạt động của Kalman Filter

Trong đó

là hiệp phương sai ước lượng (estimate covariance).

là độ lợi Kalman (Kalman Gain).

Theo như hình 4.1, chúng ta cần khởi tạo và . Vậy ta phải cung cấp 5 tham số , , , Q và R để bộ lọc Kalman hoạt động.

ta khởi tạo bằng với dữ liệu RSSI đo được lần đầu tiên.

ta sẽ cho bằng 0, sẽ tự động thay đổi sau các bước tính.

ta khởi tạo bằng với , sẽ tự động thay đổi sau các bước tính.

không có cách cài đặt cụ thể. Thông qua thử nghiệm để lấy tốt nhất.

ta lấy được bằng cách lấy dữ liệu RSSI thô và đánh giá sai số .

## Pathloss Exponent Improvement

Như ta đã biết trong chương 3, **Pathloss Exponent** là một tham số quan trọng để xây dựng phương trình RSSI Distance.

Path Loss, hay Path Attenuation, là sự giảm mật độ Path Loss có thể do nhiều tác động, chẳng hạn như mất không gian tự do, khúc xạ, nhiễu xạ, phản xạ, mất khớp nối khẩu độ-trung bình và hấp thụ. Path Loss cũng bị ảnh hưởng bởi các đường viền địa hình, môi trường (thành thị hoặc nông thôn, thảm thực vật và tán lá), môi trường lan truyền (không khí khô hoặc ẩm), khoảng cách giữa máy phát và máy thu, và chiều cao và vị trí của ăng ten. năng lượng (suy giảm) của sóng điện từ khi nó truyền qua không gian.

Ở các môi trường khác nhau, ta có thành phần Path Loss là khác nhau. Bảng 4.1và 4.2 cho ví dụ về các môi trường điển hình:

|  |  |  |
| --- | --- | --- |
| Environment | Freq (MHz) |  |
| Indoor – Retail Store | 914 | 2.2 |
| Indoor – Grocery Store | 914 | 1.8 |
| Indoor – Hard Partition Office | 1500 | 3.0 |
| Indoor – Soft Partition Office | 900 | 2.4 |
| Indoor – Factory (LOS(\*)) | 1900 | 2.6 |
| Indoor – Factory (LOS) | 1300 | 1.6-2.0 |
| Indoor – Suburban Home | 4000 | 2.1 |
| Indoor – Factory (Obstructed) | 1300 | 3.3 |
| Indoor – Factory (Obstructed) | 4000 | 2.1 |
| Indoor – Office Same Floor | 914 | 2.68-4.01 |
| Indoor – Office Entire Building | 914 | 3.54-4.33 |
| Indoor – Office Wing | 914 | 2.68-4.01 |

Table 4.1 Pathloss Exponent cho các môi trường cụ thể

|  |  |  |
| --- | --- | --- |
| Environment | Freq (MHz) |  |
| Indoor – Average | 914 | 3.14 |
| Indoor – Through One Floor | 914 | 4.19 |
| Indoor – Through Two Floor | 914 | 5.04 |
| Indoor – Through Three Floor | 914 | 5.22 |

(\*)LOS: Line of Sight

Table 4.2 Pathloss Exponent cho các môi trường cụ thể (tiếp theo)

Qua bảng trên ta có thể thấy, tham số Pathloss phụ thuộc khá nhiều vào môi trường xung quanh, ngoài ra còn có tần số của tín hiệu. Vậy nếu chúng ta đặt Pathloss Exponent là một hằng số cố định thì có thể sẽ gây ra sai số khá lớn, từ đó làm mất ổn định hệ thống, khó có thể phán đoán chính xác được vị trí thiết bị.

Vậy vấn đề cần đặt ra là: Nếu muốn độ chính xác thuật toán càng cao, thì yếu tố phụ thuộc môi trường phải giảm thiểu. Yếu tố môi trường ở thuật toán này chính là Pathloss Exponent.

Để giải quyết vấn đề trên, ta sẽ đặt thêm một thiết bị thứ hai. Thiết bị này sẽ được cố định tại một vị trí cụ thể, từ đó có thể suy ra được tham số Pathloss của môi trường.

Biến đổi phương trình (3.2) ta được:

|  |  |  |
| --- | --- | --- |
|  |  | (4.6) |

Vậy với (RSSI tại 1m) và đã được cố định sẵn thì ta chỉ cần đo được và áp dụng công thức (4.6) ta sẽ có được Pathloss Exponent và đưa tham số này vào phương trình tính khoảng cách ta sẽ có kết quả chính xác hơn từ đó tối ưu được vị trí cần tính.

## Particle Swarm Optimization

### Giới thiệu về giải thuật di truyền và tối ưu bầy đàn

Giải thuật di truyền (Genetic Algorithm – GA) là kỹ thuật phỏng theo quá trình thích nghi tiến hóa của các quần thể sinh học dựa trên học thuyết Darwin. GA là phương pháp tìm kiếm tối ưu ngẫu nhiên bằng cách mô phỏng theo sự tiến hóa của con người hay của sinh vật. Tư tưởng của thuật toán di truyền là mô phỏng các hiện tượng tự nhiên, là kế thừa và đấu tranh sinh tồn.

GA thuộc lớp các giải thuật xuất sắc nhưng lại rất khác các giải thuật ngẫu nhiên vì chúng kết hợp các phần tử tìm kiếm trực tiếp và ngẫu nhiên. Khác biệt quan trọng giữa tìm kiếm của GA và các phương pháp tìm kiếm khác là GA duy trì và xử lý một tập các lời giải, gọi là một quần thể (population).

Trong GA, việc tìm kiếm giả thuyết thích hợp được bắt đầu với một quần thể, hay một tập hợp có chọn lọc ban đầu của các giả thuyết. Các cá thể của quần thể hiện tại khởi nguồn cho quần thể thế hệ kế tiếp bằng các hoạt động lai ghép và đột biến ngẫu nhiên – được lấy mẫu sau các quá trình tiến hóa sinh học.

Ở mỗi bước, các giả thuyết trong quần thể hiện tại được ước lượng liên hệ với đại lượng thích nghi, với các giả thuyết phù hợp nhất được chọn theo xác suất là các hạt giống cho việc sản sinh thế hệ kế tiếp, gọi là cá thể (individual). Cá thể nào phát triển hơn, thích ứng hơn với môi trường sẽ tồn tại và ngược lại sẽ bị đào thải. GA có thể dò tìm thế hệ mới có độ thích nghi tốt hơn.

GA giải quyết các bài toán quy hoạch toán học thông qua các quá trình cơ bản: lai tạo (crossover), đột biến (mutation) và chọn lọc (selection) cho các cá thể trong quần thể. Dùng GA đòi hỏi phải xác định được: khởi tạo quần thể ban đầu, hàm đánh giá các lời giải theo mức độ thích nghi – hàm mục tiêu, các toán tử di truyền tạo hàm sinh sản.

Phương pháp tối ưu bầy đàn (Particale Swarm Optimization – PSO) là một dạng của thuật toán tiến hóa quần thể như giải thuật di truyền (GA). Tuy vậy PSO khác với GA ở chỗ nó thiên về sử dụng sự tương tác giữa các cá thể trong một quần thể để khám phá không gian tìm kiếm. PSO là kết quả của sự mô hình hóa việc đàn chim bay đi tìm kiếm thức ăn cho nên nó thường được xếp vào các loại thuật toán có sử dụng trí tuệ bầy đàn.

Được giới thiệu vào năm 1995 tại một hội nghị của IEEE bởi James Kennedy và kỹ sư Russell C. Eberhart. Thuật toán có nhiều ứng dụng quan trọng trong tất cả các lĩnh vực mà ở đó đòi hỏi phải giải quyết các bài toán tối ưu hóa.

Để hiểu rõ thuật toán PSO hãy xem một ví dụ đơn giản về quá trình tìm kiếm thức ăn của một đàn chim. Không gian tìm kiếm thức ăn lúc này là toàn bộ không gian ba chiều mà chúng ta đang sinh sống. Tại thời điểm bắt đầu tìm kiếm cả đàn bay theo một hướng nào đó, có thể là rất ngẫu nhiên. Tuy nhiên sau một thời gian tìm kiếm một số cá thể trong đàn bắt đầu tìm ra được nơi có chứa thức ăn. Tùy theo số lượng thức ăn vừa tìm kiếm, mà cá thể gửi tín hiệu đến các các cá thể khác đang tìm kiếm ở vùng lân cận. Tín hiệu này lan truyền trên toàn quần thể. Dựa vào thông tin nhận được mỗi cá thể sẽ điều chỉnh hướng bay và vận tốc theo hướng về nơi có nhiều thức ăn nhất. Cơ chế truyền tin như vậy thường được xem như là một kiểu hình của trí tuệ bầy đàn. Cơ chế này giúp cả đàn chim tìm ra nơi có nhiều thức ăn nhất trên không gian tìm kiếm vô cùng rộng lớn.

Như vậy đàn chim đã dùng trí tuệ, kiến thức và kinh nghiệm của cả đàn để nhanh chóng tìm ra nơi chứa thức ăn. Bây giờ chúng ta mô hình hóa mô hình sinh học này thường được gọi là quá trình phỏng sinh học mà chúng ta thường thấy trong các ngành khoa học khác. Một thuật toán được xây dựng dựa trên việc mô hình hóa các quá trình trong sinh học được gọi là thuật toán phỏng sinh học (bioinspired algorithms).

Hãy xét bài toán tối ưu của hàm số F trong không gian n chiều. Mỗi vị trí trong không gian là một điểm tọa độ n chiều. Hàm F là Hàm mục tiêu(fitness function) xác định trong không gian n chiều và nhận giá trị thực. Mục đích là tìm ra điểm cực tiểu của hàm F trong miền xác định nào đó. Ta bắt đầu xem xét sự liên hệ giữa bài toán tìm thức ăn với bài toán tìm cực tiểu của hàm theo cách như sau. Giả sử rằng số lượng thức ăn tại một vị trí tỉ lệ nghịch với giá trị của hàm F tại vị trí đó. Có nghĩa là ở một vị trí mà giá trị hàm F càng nhỏ thì số lượng thức ăn càng lớn. Việc tìm vùng chứa thức ăn nhiều nhất tương tự như việc tìm ra vùng chứa điểm cực tiểu của hàm F trên không gian tìm kiếm.

Bài toán trên chính là một bài toán PSO điển hình. Mục tiêu của chúng ta là áp dụng phương pháp tối ưu bầy đàn này vào hệ thống của chúng ta để tối ưu được sai số, tăng độ chính xác.

### Áp dụng giải thuật Particle Swarm Optimization

Có 2 yếu tố chính để xây dựng nên một bài toán Particle Swarm Optimization chính là xác định được không gian tìm kiếm của bài toán và xác định được hàm mục tiêu của bài toán đó.

Theo như đã đề cập ở chương 3, sau khi giải được phương trình (3.10) ta sẽ có được kết quả gần đúng của bài toán, nhưng kết quả này không phải là kết quả tối ưu nhất. Tuy nhiên, ta có thể sử dụng kết quả này làm trung tâm của không gian, sau đó ta tìm kiếm xung quanh điển này với một bán kính cho trước (ví dụ 5m).

Việc tiếp theo, ta sẽ tìm hàm mục tiêu (Fitness Function) phù hợp với yêu cầu bài toán đặt ra. Ta nhận thấy, bản chất của bài toán chính là tối ưu được sai số của thiết bị so với vị trị thực tế. Vật hàm mục tiêu của chúng ta chính là sai số đạt được nhỏ nhất.

Xét hệ phương trình (3.3), như ta đã biết, các phép đo này là không chính xác bởi vì chịu tác động của các yếu tố ngoại cảnh như thời tiết, vật cản,… hoặc yếu tố nội cảnh như nguồn điện không ổn định, tính chính xác của antenna,… Tuy nhiên, ta sẽ sử dụng hệ phương trình này làm thước đo sai số cho bài toán.

Giả sử trong không gian tìm kiếm, ta có phần tử thứ có vị trí , khi đó ta có phương trình khoảng cách của phần tử này đến các điểm tham chiếu như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (4.7) |

Trong đó:

là vị trí của phần tử thứ .

là vị trí của điểm tham chiếu thứ ().

Là khoảng cách của phần tử thứ k đến điểm tham chiếu thứ .

Vậy sai số khoảng cách lúc này chính là:

|  |  |  |
| --- | --- | --- |
|  |  | (4.8) |

Như vậy lúc này ta sẽ chọn hàm Fitness chính là tổng sai số của khoảng cách của phần tử đối với các vị trí tham chiếu. Tuy nhiên, theo như công thức (4.8) có thể là một số âm hoặc dương, vì vậy để nhất quán, ta sẽ chọn hàm Fitness chính là tổng bình phương sai số của khoảng cách của phần tử đối với các vị trí tham chiếu. Cuối cùng ta có hàm Fitness như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (4.9) |

Ngoài ra, một việc cũng quan trọng không kém đó chính là tổ chức quần thể. Việc tổ chức quần thể được biểu hiện qua việc khởi tạo xây dựng quần thể đó và phương pháp duy trì, phát triển quần thể.

Để khởi tạo xây dựng quần thể, ta cần biết các thông tin sau: trung tâm quần thể, bán kính không gian tìm kiếm và số lượng cá thể trong quần thể. Các thông tin này có thể được cài đặt trước bởi người dùng. Giả sử ta có trung tâm quần thể là , Bán kính không gian tìm kiếm là , số lượng cá thể là . Ta sẽ có các cá thể trong quần thể được sinh ra như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (4.10) |

Trong đó:

là vị trí của phần tử thứ .

là vị trí của trung tâm quần thể

là khoảng cách của phần tử thứ k đến trung tâm quần thể

là góc của phần tử thứ k.

là bán kính không gian quần thể.

Như vậy ta đã khởi tạo xây dựng xong quần thể. Việc tiếp theo của chúng ta là duy trì và phát triển quần thể.

Đầu tiên, để duy trì được quần thể thì mỗi cá thể sẽ được lặp đi lặp lại quá trình tìm kiếm để xác định được giải pháp tối ưu nhất. Trong mỗi vòng lặp, các phần tử sẽ tự theo dõi vị trí tốt nhất của chính nó , được xem như **Personal Best (pbest)**, giải pháp tốt ưu nhất của tất cả pbest chính là **Global Best (gbest)**.

Global Best đóng vai trò quan trọng trong việc phát triển quần thể. Để dễ hiểu, ta ví dụ như có một đàn kiến, tản ra xung quanh để tìm kiếm thức ăn. Giả sử đàn kiến có một cách liên lạc tầm xa, thì mỗi khi một con kiến tìm được khu vực có thức ăn nhiều nhất mà nó tìm thất thì nó sẽ báo về lại đàn, lúc này khu vực có thức ăn mà con kiến này tìm được chính là Personal Best. Khu vực có nhiều thức ăn nhất trong tất cả khu vực mà các con kiến báo về chính là Global Best.

Mỗi khi khu vực có nhiều thức ăn nhất trong không gian tìm kiếm được cập nhật, thì cả đàn kiến sẽ có xu hướng tiến về khu vực đó để “khai thác”. Trong lúc tiến về khu vực tốt nhất, trên đường đi có thể các con kiến sẽ kiếm được khu vực khác tốt hơn. Khi đó, Global Best được cập nhật, vậy là cả đàn có mục tiêu mới.

Quá trình lặp lại việc tìm kiếm Personal Best và Global Best chính là cơ sở để phát triển quần thể. Quá trình lặp lại có thể được biểu diễn như sau:

|  |  |  |
| --- | --- | --- |
|  |  | (4.11) |
|  |  | (4.12) |
|  |  | (4.13) |

Trong đó:

là vận tốc của cá thể thứ .

là vị trí hiện tại của phần tử thứ (thuộc không gian tìm kiếm).

là nhân tố quán tính.

là hệ số gia tốc.

là hệ số co thắt ngẫu nhiên trong phạm vi .

là số thứ tự bước lặp hiện tại và là tổng số bước lặp.

Sau mỗi bước lặp tìm kiếm, ta sử dụng hàm Fitness để cập nhật lại Personal Best cũng như Global Best. Cuối cùng, sau khi kết thúc tìm hiểm, ta sẽ tìm được vị trí tối ưu nhất cho bài toán.

Đánh giá sơ lược bài toán, giả sử ta xây dựng một quần thể có bán kính 5m với 100 cá thể lúc ban đầu, nếu lặp 50 lần, trường hợp xấu nhất sẽ tạo ra 100\*50 vị trí khác nhau. Vậy lúc này mỗi vị trí sẽ chiếm diện tích trung bình khoảng :

Lúc này sai số đạt được khoảng:

Đây là sai số phụ thuộc vào các phép đo của thiết bị. Giả sử các thiết bị đo chính xác thì sai số đạt được là 8.86cm.