



**دانشگاه پیام نور تهران**

**واحد شهر ری**

**سمینار**

**عنوان:**

**رویکرد مبتنی بر مدل برای پردازش داده ها در بستر اينترنت اشياء IoT**

**استاد راهنما:**

**جناب آقای دکتر سید علی رضوی ابراهیمی**

**نگارش:**

**مجيد لطفي**

**پاييز 1400**

چکیده

پیشرفت‌های اخیر در زمینه‌های مختلف، از جمله فناوری‌های حسگر، شبکه و پردازش داده‌ها، باعث شده اند چشم‌انداز اینترنت اشیا (IoT) هر روز بیشتر و بیشتر به واقعیت تبدیل شود. در نتیجه این پیشرفت‌ها، اینترنت اشیا امروزی امکان توسعه برنامه‌های کاربردی پیچیده برای محیط‌های اینترنت اشیا، مانند شهرهای هوشمند، خانه‌های هوشمند یا کارخانه‌های هوشمند را فراهم می‌کند و باعث شده با توجه به تبادل مکرر داده‌ها آنها به شکل جریان‌های داده دربيايد.

با این حجم فزاینده داده ای که به طور مداوم پردازش می شود، چالش های متعددی چون جلوگيري از تداخل و آسيب در فرآيند هاي گذرا به وجود می آید كه نيازمند بررسي پردازش مبتنی بر جریان داده در محیط های IoT مي باشد و از سوي ديگر با شبكه توزيع شده ناهمگون مواجه هستيم متشكل از انواع سخت افزار ها و سنسور ها كه بايد بتوان اطلاعات آنها را پردازش كرد كه بهترين محيط براي پردازش، محيط ابري مي باشد.

اما با وجود حجم بالاي اطلاعات و پردازش هاي طولاني، نمي توان تمامي اطلاعات در فضاي ابري متمركز پردازش نمود و پيشنهاد مناسب انجام اين پردازش ها در محل نزديك توليد اطلاعات مي باشد (گره پردازشي) و ارائه مدل اجراي اين ساختار بصورت جريان داده و سيستم ناهمگن و توزيع شده، هدف اصلي اين پايان نامه دكتري مي باشد.

**کلمات کلیدی:** اينترنت اشياء ، داده ، جريان داده ، پردازش ابري ، سرویس‌دهنده ابري

فهرست مطالب

[عنوان](#_Toc413125201) صفحه

# **فصل اول**

**آشنايي و معرفي**

## **مقدمه**

پیشرفت‌های اخیر در زمینه‌های مختلف، از جمله فناوری‌های حسگر، شبکه و پردازش داده‌ها، چشم‌انداز اینترنت اشیا (IoT) را قادر ساخته است که هر روز بیشتر و بیشتر به واقعیت تبدیل شود و در نتیجه این پیشرفت‌ها، اين فناوري امکان توسعه برنامه‌های کاربردی پیچیده مانند شهرها ، خانه‌ها يا کارخانه‌های هوشمند را فراهم می‌کند.

در محيط IoT با توجه به اندازه‌گیری‌های مداوم حسگر و تبادل مکرر داده‌ها بین اشیاء ، داده‌های تولید شده به شکل جریان‌های داده درآمده اند و با این حجم فزاینده داده ای که به طور مداوم پردازش می شود، چالش های متعددی براي پردازش کارآمد داده های اینترنت اشیا وجود دارد. به عنوان مثال، چگونه می توان پردازش داده های اینترنت اشیا را بدون تأثیر بر واکنش پذیری برنامه های آن تحقق بخشید. علاوه بر این، چگونه می‌توان از طریق پردازش داده‌های IoT، نیازمندی‌های مختلف عملکردی، غیرعملکردی و تعریف‌شده توسط کاربر برنامه‌ها را برآورده کرد.

در این پایان نامه دکتری، یک رویکرد کلي جدید برای پردازش برنامه های کاربردی مبتنی بر جریان داده در محیط های IoT ارائه شده است كه تمرکز آن بر قرار دادن کارآمد اپراتورهای برنامه های کاربردی در محیط های ناهمگن، توزیع شده و پویا است.

این پایان نامه دکتری توسط مدل های اطلاعاتی مختلف و تکنیک های قرار دادن اپراتور پشتیبانی می شود، به طوری که کل چرخه حیات محیط های اینترنت اشیا و برنامه های مبتنی بر جریان داده را می توان به راحتی مدیریت کرد.

در این رویکرد، یکی از اهداف اصلی پردازش داده‌های اینترنت اشیا تا حد امکان نزدیک به منابع داده است، به طوری که زیرساخت‌های ابری تنها در مواردی استفاده می‌شوند که محیط‌های اینترنت اشیا منابع پردازش کافی را برای کاربرد اینترنت اشیا ارائه نمی‌دهند.

از طریق رویکرد این پایان نامه دکتری، پردازش داده های برنامه های کاربردی اینترنت اشیا را می توان برای موارد استفاده خاص، پشتیبانی از نیازهای خاص دامنه ها، و علاوه بر این، کاربران برنامه های اینترنت اشیا، تنظیم کرد. پس از تعیین مکان‌های امکان‌پذیر، اپراتورهای پردازش با استفاده از استانداردهایی مانند TOSCA بر روی اشیاء IoT مربوطه مستقر می‌شوند و برنامه اینترنت اشیا آماده و اجرا می‌شود. در نهایت، محیط اینترنت اشیا به‌منظور شناسایی و واکنش به اختلالات مؤثر بر پردازش داده‌های برنامه‌های کاربردی اینترنت اشیا مستقر شده، به‌طور مداوم نظارت می‌شود.

رویکرد این پایان‌نامه دکترا توسط پلتفرم چند منظوره Binding and Provisioning (MBP) كه یک پلت‌فرم منبع باز اینترنت اشیا، مي باشد پشتیبانی می‌شود.

### 1-1 تعريف مسئله و سؤال

چنانچه در مقدمه بيان شد موضوع تبادل اطلاعات در محيط IoT نيازمند جريان داده و محيط ابري مي باشد و پردازش جريان داده نيز چالش خاص را داشته و نيازمند سيستم هاي توزيع شده مي باشد كه ما در اينجا با توجه به تنوع دستگاه ها، با يك سيستم ناهمگن روبرو هستيم. در اين نوع سيستم ها با توجه به حجم بالاي اطلاعات بهترين راهكار براي پردازش اطلاعات، قراردادن پردازش در اپراتورهاي پردازش (گره هاي پردازشي) نزديك به محل ايجاد اطلاعات و ارسال پردازش نهايي به سيستم ابري است. لذا سؤالاتی كه تا پايان اين تحقيق پاسخ داده می‌شود را این‌گونه بيان می‌کنیم.

1. چگونه بايد پردازش مناسب براي انواع جريان داده را تشخيص داد؟
2. چگونه مي توان اپراتورهاي پردازش جريان داده را در محيط هاي ناهمگن و پويا اجرا كرد؟
3. چگونه مي توان مدل مناسب را طراحي، اجرا و كارائي اين مدل را تضمين كرد؟

### 1-2 ضرورت تحقيق

ما در اينترنت اشياء با تنوع اشياء و داده هاي آنها مواجه هستيم كه در يك جريان داده در حال اجرا بوده و هر دستگاه و شيء در اين شبكه نيازمند پردازش هايي مي باشد كه بايد پردازش داده ها به موقع و کارآمد باشد . اين موضوع نيازمند بررسي تمام جوانب و ايجاد مدل مناسب مي باشد كه بتواند الزامات IoT را پشتيباني كرده و مدلي مناسب و آسان براي کاربران باشد.

### 1-3 اهداف

با توجه به موضوعات مطرح شده در مقدمه و ضروريات تحقيق مي توان اين اهداف را براي اين پايان نامه متصور بود:

الف) پردازش به موقع و کارآمد داده ها در محیط های اینترنت اشیاء.

ب) مدل سازی مناسب و آسان براي کاربران در محیط های اینترنت اشیاء.

ج) مدل‌سازی پردازش جریان داده و تشخيص اختلالات که الزامات اینترنت اشیا را پشتیبانی کند.

د) قرار دادن كارآمد اپراتورهای پردازش بر اساس نیازها در محیط های پويا و ناهمگن اینترنت اشیاء.

ايجاد چنين مدل هايي با دو روش خودكار و دستي امكان پذير بوده كه در اين پايان نامه به روش دستي كه در آن خود تحليل گر با توجه به الزامات و ساختار IoT تصميم مي گيرد پردازش ها و نوع ارتباط داده ها به چه صورت باشد.

### 1-4 جمع‌بندی

اینترنت اشیاء به زبان ساده، ارتباط حسگرها و دستگاه‌ها با شبکه‌‌ای است که از طریق آن می‌توانند با یکدیگر و با کاربرانشان تعامل کنند. هر كدام از اين دستگاها يا اشياء داراي داده هايي مي باشند كه مداوم و آنلاين در حال ساخت مي باشد كه خود باعث ايجاد يك جريان داده مي شود. در اين محيط از يك سو جريان داده و از سوي ديگر پردازش اين داده ها را داريم لذا بايد بدانيم اين ساختار را چگونه و با چه مدلي اجرا نماييم.

این مدل بايد منابع داده، سینک های داده، اپراتورهای پردازش و جریان داده را در میان آنها توصیف کند و علاوه بر این، الزامات اپراتورهای پردازش برای اشیاء IoT را شرح دهد.

# **فصل دوم**

**پيشينه و سابقه**

## 2-1 اينترنت اشياء

اصطلاح اینترنت اشیا (IoT) برای اولین بار در اواخر دهه 90 ظاهر شد، با ایده اشتون که به رایانه ها اجازه می دهد همه چیز را در مورد چیزها بدانند، اين ایده ابتداعا تقویت رایانه‌هایی با فناوری‌های شناسایی فرکانس رادیویی (RFID) و حسگر برای جمع‌آوری اطلاعات، مشاهده و شناسایی یک محیط بدون نیاز به کمک انسانی بود. به این ترتیب، امکان ردیابی و نظارت بر موارد به منظور کاهش هزینه ها و علاوه بر آن، اطلاع از زمان نیاز به تعمیر یا تعویض وجود دارد. ورمسان و همکاران اینترنت اشیا را به عنوان الگویی تعریف می کند که در آن چیزهای مختلفی وجود دارد.

در این محیط ، این اشياء به صورت بی سیم یا سیمی متصل می شوند، به طور منحصر به فرد قابل شناسایی هستند و می توانند برای رسیدن به اهداف مشترک با یکدیگر همکاری کنند. برای امکان پذیر ساختن این پارادایم، اینترنت اشیا از چندین فناوری توانمند منشأ گرفته از زمینه های تحقیقاتی مختلف، مانند ارتباطات ماشین به ماشین، RFID، شبکه های حسگر بی سیم (WSN)، داده های معنایی، رایانش ابری، و خدمات بهره می برد.

امروزه، بسیاری از برنامه‌های کاربردی برای اینترنت اشیا در حوزه‌های مختلفی مانند مراقبت‌های بهداشتی نظارت بر محیط زیست، یا کارخانه‌های هوشمند و حتي شهر هاي هوشمند توسعه یافته‌اند.

## 2-2 پردازش جريان داده و داده پيچيده

الزامات برنامه های کاربردی اینترنت اشیا نمی توانند با به کارگیری سیستم های مدیریت پایگاه داده سنتی (DBMS) كه با استفاده از فرآیندهای استخراج-تبدیل بار (ETL)، که نیاز به ذخیره و فهرست بندی داده ها برای پردازش دارند، برآورده شوند. بنابر اين یکی از نیازهای مهم برنامه های اینترنت اشیا، توانایی پردازش جریان‌های داده‌ای چندگانه، پیوسته، سریع و متغیر مي باشد که اين موضوع برنامه های اینترنت اشیا را قادر می سازد مقیاس پذیر، پویا و واکنش پذیر باشند.

پردازش جریان داده، که تکاملی از پردازش داده در DBMS است، پرس‌وجوهای پیوسته را بر روی جریان‌های داده ورودی اجرا می‌کند و سیستم های مدیریت جریان داده (DSMS) بر روی داده های گذرا کار میکنند، یعنی داده هایی که به طور مداوم به روز می شوند.

علاوه بر این، پرس و جوها در DBMS ها یک بار اجرا می شوند و پاسخ های کامل را برمی گردانند، در حالی که DSMS ها پرس و جوها را به طور مداوم اجرا می کنند و پس از رسیدن داده های جدید، پاسخ های به روز ارائه می دهند.

به طور معمول، DSMS جریان‌های داده را از طریق دنباله‌ای از تبدیل‌های مبتنی بر عملگرهای SQL، مانند انتخاب، تجمیع، یا پیوستن، که توسط جبر رابطه‌ای تعریف می‌شود، پردازش می‌کند. از سوی دیگر، پردازش رویدادهای پیچیده (CEP) شامل مجموعه‌ای از اصول و تکنیک‌ها برای تجزیه و تحلیل مجموعه‌هایی از رویدادها است که تا حدی بر اساس زمان با رسیدن این رویدادها مرتب شده‌اند. یعنی CEP ابزاری را برای پردازش مجموعه ای از رویدادهای مرتبط به هم به صورت پیوسته و به موقع فراهم می کند.

رویدادها معمولاً برای پردازش در الگوها ارائه می‌شوند، با این حال، می‌توانند با رویدادهای نامرتبط دیگر مخلوط شوند. ویژگی مهم CEP توانایی تشخیص الگوها (به عنوان مثال، روابط) بین رویدادها است به عنوان مثال، موقعیت‌هایي مانند خرابی ماشین. تصمیم گیری که کدام رویکرد پردازش باید استفاده شود. از سوی دیگر، اگر برنامه نیاز به تجزیه و تحلیل مجموعه ای از رویدادهای نامرتب داشته باشد، CEP باید اعمال شود.

## 2-3 مشكل قراردادن اپراتورها

پردازش جریان های داده می تواند از طریق یک سيستم متمرکز انجام شود یا می توان آن را بین گره های پردازشی مختلف برای اجرا توزیع کرد. با این حال، پردازش توزیع شده علاوه بر پردازش به موقع، چالش دیگری را نیز به همراه دارد که آن مشکل قرار دادن اپراتور مي باشد.

هدف این مشکل یافتن یک مکان بهینه از کل پرس و جوهای پیوسته یا اپراتورهای منفرد در مجموعه ای از گره های پردازشی مختلف است که در سراسر یک شبکه توزیع شده اند. مکان بهینه معمولاً بر اساس توابع هزینه تعریف شده توسط سیستم یا تعریف شده توسط کاربر محاسبه می شود، که هدف آن ارائه، برای مثال، عملکرد بالاتر یا توزیع بار بهتر است.

## 2-4 تكنولوژي TOSCA

فناوري رایانش ابری اخیراً برای میزبانی و ارائه خدمات از طریق اینترنت پدید آمده و به طور فزاینده ای همراه با پارادایم اینترنت اشیا به کار گرفته شده است تا بتواند باعث مقیاس پذیری و قابلیت همکاری آن شود. رایانش ابری می‌تواند در عين هزینه‌های پایین برای استقرار کل محیط‌های اینترنت اشیا ، قابليت راه‌اندازی و ادغام سریع اشیاء جدید و برنامه‌های کاربردی اینترنت اشیا را ايجاد کند.

فناوري TOSCA (Topology and Orchestration Specification for Cloud Applications) یک استاندارد OASIS تایید شده برای مدل سازی، استقرار و مدیریت برنامه های کاربردی ابری است كه داراي دو بخش اصلی است: (1) مدل توپولوژی یک برنامه کاربردی كه شامل اجزای نرم افزار ، پلت فرم و زیرساخت آن مي باشد، و (2) مراحل استقرار این برنامه.

TOSCA بسیار عمومی است به طوری که تعریف انواع دلخواه را برای توصیف مؤلفه های برنامه، به نام انواع گره، و وابستگی های آنها، به نام انواع رابطه، قادر می سازد. تصوير زير يك نمونه از مدل TOSCA براي يك نرم افزار كاربردي انتشار و اشتراك اطلاعات مي باشد.

****

شکل 1

TOSCA استقرار و مدیریت برنامه های ابری را در سراسر چرخه زندگیشان آسان تر می کند بدون اینکه خللی در الزامات امنیتی ، حاکمیتی و انطباقی آنها پیش آید . همچنين توانایی آن برای تسهیل یک اکوسیستم است که اجرای پورتابل برنامه ها برای ابر و بین ابرها را فراهم می سازد و این امکان را فراهم می سازد که برنامه های ابری در هر ابری به صورت پورتابل توصیف ، مدل سازی ، پکیج بندی ، تنظیم و مانیتور شوند .

# **فصل ششم**

## **مقدمه**

بحث

.

### **6-1 بحث و پیشنهاد‌ها:**

.

**6-1-1 بحث** :

می‌شود .

**6-1-2 پيشنهادها** :

هست.

### **6-2** نتیجه‌گیری :

رایانش

.

# **منابع و مراجع**

**فهرست منابع غیرفارسی و اینترنتی**

1. R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, “CryptDB: Processing queries on an encrypted database.” Commun. ACM, vol. 55, no. 9, pp. 103–111, 2012.

**ABSTRACT**

tion.