

A Study on Intelligent Design and Control for Mobile Edge Computing Systems under the Wireless Environment

著者	Koketsu Rodrigues Tiago
number	64
学位授与機関	Tohoku University
学位授与番号	情博第726号
URL	http://hdl.handle.net/10097/00130217

氏名	チアゴ コウケツ ロドリゲス TIAGO KOKETSU RODRIGUES
学位の種類	博士(情報科学)
学位記番号	情博第726号
学位授与年月日	令和2年3月25日
学位授与の要件	学位規則第4条第1項該当
研究科、専攻	東北大学大学院情報科学研究科(博士課程)応用情報科学専攻
学位論文題目	A Study on Intelligent Design and Control for Mobile Edge Computing Systems under the Wireless Environment (無線環境下におけるモバイルエッジコンピューティングシステムの知能設計と制御に関する研究)
論文審査委員	(主査) 東北大学教授 加藤 寧 東北大学教授 菅沼 拓夫 東北大学教授 西山 大樹 東北大学准教授 川本 雄一

論文内容の要旨

As much as cloud computing is useful and can enhance the capacities of user equipment by connecting them to powerful cloud servers, it has its limitations because it forces users to connect to faraway servers. This can be overcome through the edge cloud, where there are more, less powerful servers spread around the edge of the network, near the users, for them to connect to. This thesis proposes several solutions to configure the edge cloud system in order to optimize its operation. In this research, we analyze the delay to access the edge cloud and how to maximize the number of users that can connect simultaneously. The solutions will make use of machine learning to remain efficient even if the problem is complex.

The thesis first explains the background of the study. The benefits of edge cloud and the issues brought by the massive number of users and the wireless environment are fully described. This is done so even those unfamiliar with the field can understand the importance of the purpose of the study: delivering fast and cheap edge cloud service.

Then, the thesis presents all relevant existing research in edge cloud. From this group of research, the thesis presents a general service model and an equivalent completely novel mathematical model. The mathematical model considers user mobility and the full-service delay of edge cloud. It can be applied to any edge cloud research, making it a very useful tool.

Then a protocol for allocating resources to edge cloud users so that their service delay is minimized is shown. The solution indicates which server the users should connect to. The solution also determines the transmission power level of the base station. The performance of the protocol is better than existing solutions in the literature in allocating resources and lowering delay.

After that, the thesis provides a policy for deciding how to optimally deploy edge cloud servers so that the underlying service delay is minimized. The deployment is based on the location of users and the network topology of the base stations. The proposed policy is shown to support over twice as many users and require fewer resources than solutions in the literature.

Finally, a protocol for maximizing user capacity per edge cloud server while constantly respecting a maximum service delay limit is presented. The solution uses artificial intelligence for determining which server the users should connect to and how to set up base stations. The solution modifies the system as users move around can double user capacity of the edge cloud. In summary, this thesis proposes several machine learning-based methods for improving operation of the edge cloud by lowering the service delay and increasing the user capacity. Additionally, a new mathematical model of the edge cloud service is provided.

論文審査結果の要旨

クラウドコンピューティングは、ユーザの機器を強力なクラウドサーバに接続することで機器の性能を高めることができる非常に有効なものである。しかし、クラウドコンピューティングでは遠くのサーバに接続する必要があり、その機能には限界がある。この問題はエッジクラウドを利用することで克服できる。エッジクラウドではより強力なネットワークサーバがユーザの近くのネットワークエッジに配置され、ユーザとの接続を可能とする。本論文では、エッジクラウドシステムを最適に機能させるための手法を提案し、エッジクラウドへのアクセスの遅延を分析することで同時に接続できるユーザ数を最大化する手法について研究を行った。本手法では、問題が複雑な場合でもその効率を維持できるよう機械学習を利用する。この論文は6つの章で構成されている。

第1章では、この研究の背景について説明している。エッジクラウドの利点と、大量のユーザとワイヤレス環境によってもたらされる問題について説明している。また、この研究の目的と内容について述べている。

第2章では、エッジクラウドに関連する研究を紹介し、既存研究における一般的なサービスモデルおよび関連する数学モデルについて言及している。この数学モデルは、ユーザモビリティとエッジクラウドのあらゆるサービス遅延を考慮しており、エッジクラウドの研究に適用することが可能であることを述べている。

第3章では、サービスの遅延が最小となるようにエッジクラウドユーザに最適にリソースを割り当てるプロトコルについて提案している。本プロトコルは、ユーザが接続するサーバを選択するとともに、基地局の送信電力レベルを決定する。このプロトコルのパフォーマンスが、リソースの割り当てと遅延の低減に関して既存の手法よりも優れていることを示している。

第4章では、サービスの遅延を最小限に抑えるため、実行するエッジクラウドサーバを最適に選択する手法について提案している。また、提案手法は既存の手法よりも少ないリソースで2倍以上のユーザを収容することができることを示し、通信とサービスの融合という観点で貢献は大きい。

第5章では、最大サービス遅延制限を常に考慮しながら、エッジクラウドサーバあたりのユーザ容量を最大化するためのプロトコルについて述べている。このプロトコルでは、人工知能を使用して、ユーザが接続するサーバと基地局のセットアップ方法を決定する。本手法は、ユーザの使用状況に応じてシステムを変更し、エッジクラウドの収容ユーザ数を2倍にすることを可能とし、極めて評価の高い提案である。

第6章は、本論文の要約である。

以上のように、本論文では、サービス遅延を減らし、収容ユーザ数を向上させることにより、エッジクラウドの動作を改善するためのいくつかの機械学習ベースの方法を提案し実験によりその有用性を実証した。

よって、本論文は博士（情報科学）の学位論文として合格と認める。