RGC Ref. No.:
UGC/FDS16/E09/21
(please insert ref. above)

RESEARCH GRANTS COUNCIL
COMPETITIVE RESEARCH FUNDING SCHEMES FOR
THE LOCAL SELF-FINANCING DEGREE SECTOR

FACULTY DEVELOPMENT SCHEME (FDS)

Completion Report
(for completed projects only)

Submission Deadlines:
1. *Auditor's report with unspent balance, if any: within **six** months of the approved project completion date.*
2. *Completion report: within **12** months of the approved project completion date.*

Part A: The Project and Investigator(s)

1. Project Title

An Investigation of Reliability-Aware Wireless Edge Caching Networks (WECNs) with Bundle Recommendation

2. Investigator(s) and Academic Department(s) / Unit(s) Involved

Research Team	Name / Post	Unit / Department / Institution
Principal Investigator	Dr. Yaru FU / Assistant Professor	Department of Electronic Engineering and Computer Science, School of Science and Technology, Hong Kong Metropolitan University
Co-Investigator	Dr QUEK Tony Quee-seng / Professor	Information Systems Technology and Design, Singapore University of Technology and Design
Co-Investigator	Prof. WONG Angus Kin-yung / Professor	(Formerly) Department of Electronic Engineering and Computer Science, School of Science and Technology, Hong Kong Metropolitan University

3. Project Duration

	Original	**Revised**	**Date of RGC / Institution Approval** *(must be quoted)*
Project Start Date	01/01/2022	N/A	N/A
Project Completion Date	31/12/2023	N/A	N/A
Duration *(in month)*	24	N/A	N/A
Deadline for Submission of Completion Report	31/12/2024	N/A	N/A

4.4 Please attach photo(s) of acknowledgement of RGC-funded facilities / equipment.
N/A

Part B: The Final Report

5. Project Objectives

5.1 Objectives as per original application

1. To design appropriate fault-tolerant coding schemes to ensure the reliability of the WECNs (wireless edge caching networks), namely, the users can successfully retrieve their desired content items even though the cached packets of each content in multiple edge nodes (ENs) are corrupted at the same time, and besides the least cost incurred by repairing failed nodes and the load balance among ENs are achieved.

2. To design bundle algorithms for recommendation-aware WECNs to recommend contents from a widened and flexible horizon for individual users by mathematically analyzing the effect of personalized bundle recommendations on users' bundle preference distribution and content consumption pattern.

3. To investigate the impact of the interplay between bundle recommendation and cache placement on WECN's effectiveness, such as achievable revenue and cache hit ratio, by designing joint personalized bundle recommendation and cache decision approaches, taking into account the constraints of successful content retrieval and repair, total repair cost requirement, cache capacity budget of each EN, and per user's recommendation quality and quantity.

5.2 Revised objectives

Date of approval from the RGC: _____

Reasons for the change: _____

_____

1.

2.

3. ..

5.3 Realisation of the objectives

(Maximum 1 page; please state how and to what extent the project objectives have been achieved; give reasons for under-achievements and outline attempts to overcome problems, if any)

(1) Realisation of Objective 1:

Regarding the reliability consideration for wireless content caching networks, we utilized a systematic erasure code to ensure the reliability of the system. Building upon this, we investigated the problem of minimizing costs in wireless content caching networks by optimizing the cache placement decisions. The total cost is comprised of three primary components: caching cost, retrieving cost, and update cost (due to the applied coding scheme). To tackle the non-convex integer programming problem, we developed both optimal and suboptimal solutions that strike a balance between time efficiency and system performance.

(2) Realisation of Objectives 2 and 3:

Regarding the cache efficiency of wireless content caching networks, we proposed the utilization of bundle recommendation. However, solving the problem driven by bundle recommendation is more complex than the single-item recommendation problem. This complexity arises from the need to address at least three key issues: 1) determining the total number of bundles per user and the specific contents for each bundle; 2) recommending bundles to different users, considering their diverse preferences; and 3) deciding the cache placement at the edge nodes. To address this challenging problem, we developed three distinct joint bundle recommendation and content placement decision-making algorithms based on alternative optimization theory.

Apart from demonstrating the benefits of the interplay between recommendation and caching in improving the performance wireless content caching systems, we extensively investigated the joint caching, recommendation, and routing problem. As it is essential not to overlook the impact of the routing strategy on cache placement strategies. To this end, we thoroughly characterized the content retrieval process for each user and formulated a cost minimization problem that considers practical constraints such as cache capacity budgets per user and base station, as well as recommendation quality and quantity requirements. To handle the complexity of the problem, we divided it into two sub-problems. We addressed these sub-problems individually and iteratively optimized them in an alternating fashion until convergence was achieved.

(3) Extension of the objectives:

Later on, we recognized that relying solely on cache hit ratio (CHR) may not provide a complete evaluation of content caching performance. It’s important to note that solely pursuing a high CHR can result in recommended items that are too similar to each other. This limits users’ choices and can potentially reduce the long-term revenue of the system. In this context, we made the first attempt to involve the effect of recommendation diversity in cache-enabled wireless cellular networks, aiming to maximize the utility of the system. We rigorously proved the NP-hardness of the maximization problem and solved it in an efficient manner. More specifically, we utilized the dynamic programming algorithm and propose an efficient decision-making strategy to determine the cache placement at the base station and the personalized recommendation.

5.4 Summary of objectives addressed to date

Objectives *(as per 5.1/5.2 above)*	**Addressed** *(please tick)*	**Percentage Achieved** *(please estimate)*
1. To design appropriate fault-tolerant coding schemes to ensure the reliability of the WECNs (wireless edge caching networks), namely, the users can successfully retrieve their desired content items even though the cached packets of each content in multiple edge nodes (ENs) are corrupted at the same time, and besides the least cost incurred by repairing failed nodes and the load balance among ENs are achieved.	✓	100%
2. To design bundle algorithms for recommendation-aware WECNs to recommend contents from a widened and flexible horizon for individual users by mathematically analyzing the effect of personalized bundle recommendations on users' bundle preference distribution and content consumption pattern.	✓	100%
3. To investigate the impact of the interplay between bundle recommendation and cache placement on WECN's effectiveness, such as achievable revenue and cache hit ratio, by designing joint personalized bundle recommendation and cache decision approaches, taking into account the constraints of successful content retrieval and repair, total repair cost requirement, cache capacity budget of each EN, and per user's recommendation quality and quantity.	✓	100%

6. Research Outcome

6.1 Major findings and research outcome
(Maximum 1 page; please make reference to Part C where necessary)

In [1], we utilized a systematic erasure code in wireless content caching systems to ensure the reliability of the system. Building upon this, we investigated the problem of minimizing costs in wireless content caching networks by periodically optimizing the cache placement decisions. The total cost is comprised of three primary components: caching cost, retrieving cost, and update cost. To tackle the non-convex integer programming problem, we developed both optimal and suboptimal solutions that strike a balance between time efficiency and system performance. Extensive numerical results demonstrate that our suboptimal solution achieves near-optimal performance and surpasses all benchmark strategies.

In [2], we discussed the current research trends, challenges, and potential solutions for wireless content caching systems. Therein, we proposed the utilization of bundle recommendation to enhance the cache efficiency of content caching networks. However, solving the problem driven by bundle recommendation is more complex than the single-item recommendation problem. This complexity arises from the need to address at least three key issues: 1) determining the total number of bundles per user and the specific contents for each bundle; 2) recommending bundles to different users, considering their diverse preferences; and 3) deciding the cache placement at the base stations. To tackle this challenging problem, we developed three distinct joint bundle recommendation and content placement decision-making algorithms based on alternative optimization theory. We provide a thorough examination of the computational requirements and convergence properties of the proposed algorithms to ensure their practical applicability and efficiency. Comprehensive numerical simulations demonstrate that bundle recommendation can achieve a higher cache hit ratio and greater system revenue compared to conventional single-item recommendation schemes. The underlying reason is that bundle recommendation can further influence or nudge the preferences of users, encouraging them to request similar content items. These results highlight the effectiveness of bundle recommendation in improving the cache efficiency of wireless content caching networks.

Despite demonstrating the benefits of the interplay between caching and recommendation in improving the performance of pure caching schemes [1-2], it is crucial not to overlook the impact of the routing strategy on cache placement strategies. In [3], we extensively investigated the joint caching, recommendation, and routing problem in wireless edge caching systems. We thoroughly characterized the content retrieval process for each user and formulated a cost minimization problem that considers practical constraints such as cache capacity budgets per user and base station, as well as recommendation quality and quantity requirements. To handle the complexity of the problem, we divided it into two sub-problems: the recommendation design problem and the joint caching and routing optimization problem. We addressed these sub-problems individually and iteratively optimized them in an alternating fashion until convergence was achieved. Specifically, we employed the simulated annealing algorithm to effectively solve the recommendation decision-making problem. For the joint caching and routing design sub-problem, we transformed binary variables into continuous ones through relaxation and utilized the alternating direction method of multipliers algorithm to solve the relaxed problem. Extensive numerical results not only confirmed the convergence performance of our proposed joint optimization algorithm but also demonstrated its superiority over various benchmark schemes. This superiority was evident across different network metrics, including overall cost and cache hit ratio. In [4], we made the first attempt to involve the effect of recommendation diversity in cache-enabled wireless cellular networks, aiming to maximize the utility of the system. A two-fold method was developed to alternately optimize these two types of binary variables until convergence. We performed extensive numerical simulations to notarize the convergence performance of our designed joint optimization algorithm and its superiority in terms of utility, diversity, and CHR compared to various baseline schemes considering or without considering diversified recommendation.

6.2 Potential for further development of the research and the proposed course of action *(Maximum half a page)*

In this project (UGC/FDS16/E09/21), the main objectives were to enhance the reliability (addressing data corruption issues) and cache efficacy (improving cache hit ratio) of wireless edge caching networks. To tackle these challenges, the project proposed fault-tolerant codes and personalized bundle recommendation algorithms. The focus was on caching content files, without considering collaborative designs for communication, execution, and caching. The methodologies employed in this project were combinatorial optimization and dynamic programming.

Moving forward, our further development aims to study high-effectiveness service-aware wireless edge caching systems for the intelligent 6G era. This research encompasses various aspects, starting from the analysis of user demand modeling to the joint recommendation decision-making and service placement optimization algorithms. In service-aware wireless edge caching systems, lightweight and independent micro-services (MS) are combined to enable different large-scale and intelligent applications or services. The successful execution of each service relies on the successful computation of individual component MSs, and multiple duplications of each MS can be installed to handle simultaneous requests.

Within this context, we collectively investigate the convergence of communication, execution, and caching in service-aware wireless edge caching systems, with the aim of providing enhanced achievable revenue and high-quality data services for network operators and the general public. In designing the corresponding solutions, different models including revised multinomial logit choice models, optimization models, reinforcement learning models, and efficient neural network models are applied.

7. Layman's Summary

(Describe <u>in layman's language</u> the nature, significance and value of the research project, in no more than 200 words)

The rapid growth of mobile networks has led to the emergence of intelligent applications such as virtual reality, online games, and high-definition video services. These applications require fast response times and handle large amounts of data. However, accessing content from remote servers can be challenging due to network congestion and limited capacity. To address these challenges, wireless edge caching has been proposed as a solution. It involves storing popular and reusable content at edge nodes like base stations, access points, and mobile devices. In this project, our aim was to address two key issues related to wireless edge caching: reliability and cache effectiveness. For reliability, we developed a coding scheme that allows users to repair and retrieve desired content even if the cached packets in multiple edge nodes are corrupted simultaneously. To improve cache effectiveness, we designed algorithms that recommend content bundles to users and maximize network efficiency by considering factors like content retrieval, repair costs, and cache capacity. These solutions offer benefits to both networks and end users. They ensure network reliability, reducing maintenance costs for networks, and provide subscribers with a better data service experience.

Part C: Research Output

8. Peer-Reviewed Journal Publication(s) Arising Directly From This Research Project

(Please attach a copy of the publication and/or the letter of acceptance if not yet submitted in the previous progress report(s). All listed publications must acknowledge RGC's funding support by quoting the specific grant reference.)

The Latest Status of Publications				**Author(s)** *(denote the corresponding author with an asterisk*)*	**Title and Journal / Book** *(with the volume, pages and other necessary publishing details specified)*	**Submitted to RGC** *(indicate the year ending of the relevant progress report)*	**Attached to this Report** *(Yes or No)*	**Acknowledged the Support of RGC** *(Yes or No)*	**Accessible from the Institutional Repository** *(Yes or No)*
Year of Publication	**Year of Acceptance** *(For paper accepted but not yet published)*	**Under Review**	**Under Preparation** *(optional)*						
2022				Yaru Fu, Jianqing Liu*, Junming Ke, John Kowk Tai Chui, Kevin King Fai Hung	Optimal and Suboptimal Dynamic Cache Update Algorithms for Wireless Cellular Networks, *IEEE Wireless Communications Letters*, vol. 11, no. 12, pp. 2610-2614, Dec. 2022.	Yes (12/2022) Reference [1] in Section 6.1.	Yes Attachment [1]	Yes	Yes
	2023			Yaru Fu*, Yue Zhang, Qi Zhu, Hong-ning Dai, Mingmei Li, and Tony Q. S. Quek	A New Vision of Wireless Edge Caching Networks (WECNs): Issues, Technologies, and Open Research Trends, Early access, *IEEE Network*, 2023. DOI: 10.1109/MNET.124.2200003.	No Reference [2] in Section 6.1.	Yes Attachment [2]	Yes	Yes

2024				Yu Hua, Yaru Fu*, and Qi Zhu	On the design of cost minimization for D2D-enabled wireless caching networks: A joint recommendation, caching, and routing perspective, *IET Communications*, https://doi.org/10.1049/cmu2.12716, pp. 129-144, Jan. 2024	No Reference [3] in Section 6.1.	Yes Attachment [3]	Yes	Yes
	2023			Yaru Fu, Yue Zhang*, Zheng Shi, Hong Wang, and Quan Yu	Utility Maximization for Wireless Content Caching Networks with Diversified Recommendation, *IEEE Transactions on Vehicular Technology*, 2023, DOI: 10.1109/TVT.2023.3339755	No Reference [4] in Section 6.1.	Yes Attachment [4]	Yes	Yes

9. Recognized International Conference(s) In Which Paper(s) Related To This Research Project Was / Were Delivered
(Please attach a copy of each conference abstract)

Month / Year / Place	**Title**	**Conference Name**	**Submitted to RGC** *(indicate the year ending of the relevant progress report)*	**Attached to this Report** *(Yes or No)*	**Acknowledged the Support of RGC** *(Yes or No)*	**Accessible from the Institutional Repository** *(Yes or No)*
08/2023/ Dalian, China	Energy Consumption Minimization for Distributed Microservice-Aware Wireless Cellular Networks	The 12th IEEE/CIC International Conference on Communications in China (ICCC 2023)	No	Yes Attachment [5]	Yes	Yes
11/2023/Hangzhou, China	Joint Resource Allocation and Cache Placement for RSMA-Aware Multi-BS Cooperative Networks	2023 International Conference on Wireless Communications and Signal Processing (WCSP)	No	Yes Attachment [6]	Yes	Yes

10. Whether Research Experience And New Knowledge Has Been Transferred / Has Contributed To Teaching And Learning
(Please elaborate)

A new chapter regarding the studied topics of this project was involved by the course ELEC S306F (Wireless Networks, 2022 Autumn Semester), wherein the PI, Dr. Yaru Fu, is the course coordinator. This new chapter introduces the latest technologies of wireless cellular networks, as proposed by this project.

11. Student(s) Trained
(Please attach a copy of the title page of the thesis)

Name	**Degree Registered for**	**Date of Registration**	**Date of Thesis Submission / Graduation**
N/A			

12. Other Impact
(e.g. award of patents or prizes, collaboration with other research institutions, technology transfer, teaching enhancement, etc.)

1. An invited talk with regard to the studied topics of this project was given to China University of Petroleum, hosted by Dr. Zhanglei Shi, on 30/09/2022.

2. A similar talk was delivered to Guangdong University of Technology, hosted by Prof. Yongquan Ling.

13. Statistics on Research Outputs

	Peer-reviewed Journal Publications	Conference Papers	Scholarly Books, Monographs and Chapters	Patents Awarded	Other Research Outputs (please specify)	
No. of outputs arising directly from this research project	4	2			Type	No.

14. Public Access Of Completion Report

(Please specify the information, if any, that cannot be provided for public access and give the reasons.)

Information that Cannot Be Provided for Public Access	Reasons