

# 5G 网络快取与传输效能改善 ——以 Facebook（脸书）内容递送为例

蔡汶儒，温演福 台北大学信息管理研究所

**摘要** 随着 5G（第五代移动通信技术）网络即将到来，互联网设备、无线网络流量和基地台数量都将会增加，基地台的架设密度也将会变得密集。5G 网络的速度变得更快，传输也会变得更拥挤，若可以在基地台里面暂时存放常被存取的资料，即可减少回程的负载，但基地台里的快取空间有限，不太可能容纳得下所有装置可能会请求的资料，如何在有限的空间摆放快取资料提升装置请求资料的命中率成为一个重要的议题。许多诸如 Facebook 的动态网络资料随时间而存取不同的网页，但是网页是由主机决定，许多使用者的 Web Part（如广告页）是相同的，在这种情境之下，本研究将进一步探讨如何运用 5G 边缘快取，以提升系统存取效能。

**关键词** 5G；快取储存；资源分配；雾运算

## 5G Cache and Transmission Performance Improvement— The Case of Facebook Content Delivery

Wen-Ju Tsai, Yean-Fu Wen

The Institute of Management Information System, Taipei University

**Abstract** With the arrival of the 5G network, the number of Internet devices, wireless network traffic and base stations will increase, and the density of the base stations will become denser. 5G networks will become faster and the traffic will become congested. If the frequently accessed data can be temporarily stored in the base station, the backhaul load can be reduced, but the cache space in the base station is limited and it is unlikely to accommodate all the information that may be requested by the device. How to place cache data in limited space to enhance the hit rate of data request has become an important issue. The dynamical Web data are changed time by time, but all these data are determined by server hosts. Some Web part, such as advertising pages, might be the same among the users. This study further discusses how to use 5G edge caching to enhance the system access performance.

**Keywords** 5G; Cache storage; Resource allocation; Fog Computing

## 1 绪论

### 1.1 研究背景

5G 网络能提供极快的传输速度,但是它属于极高频,特性是传输距离比较短,穿墙能力有限,因此需要通过兴建大量密集的基地台覆盖以解决传输距离较短这个问题。

装置发出请求的资料量往往大于基地台里面的储存空间,在储存空间有限的情况下,基地台内只能选择快取装置最有可能会去存取的资料,并因应不同装置的连线尽可能提升快取资料的命中率而不定期地替换资料。

装置使用者热衷于下载短暂的内容,替换快取通常会使用 LRU (least recently used, 最近使用过的) 方式,意思是让最近被使用的资料替换掉最久没被使用的资料,但是因为基地台每天收到的资料请求数量不一定,很难判定资料是否之后还会被请求,因此这样的方法在本研究中并不适用。改善 LRU 缺点的解决方法可以使用 LRU with Prefilters (LRU 带前置滤波器),事先预测内容的 LRU<sup>[1]</sup>,这样的方式有点类似推荐系统,根据装置已知的资料(如装置发出资料请求的记录),预测什么类型的资料会容易被装置存取,但这个方法和推荐系统有一点不同,推荐系统是根据使用者的历史浏览记录和购买记录推荐等已知的资料推荐使用者可能感兴趣的资料,但是使用者是否要存取是由装置自行决定,装置需要主动去接收资料,系统所推荐的资料不一定会被存取。这里的方法也是根据对装置已知的信息,做出决策,将装置可能会存取的资料内容事先存

在基地台内的快取中<sup>[2]</sup>,装置会对基地台发出请求,被动接收已事先存在快取中的资料,装置是被动接收资料。

装置上的应用程序较多,种类也不一样,要找出一个包容所有种类应用程序的基地台布署方式并不容易,因此设定单一应用程序为例研究,找出较佳基地台资料布署方式。

如图 1 所示,2016 年 Facebook 在台湾使用率约为 82%<sup>[3]</sup>,使用率很高,是比较常见的社交软件,Facebook 涵盖了 Line (连我) 和 Instagram (照片墙) 的主要功能 (Line 主要功能为信息传送的作用,Instagram 主要功能为图片动态的作用),因此选择 Facebook 为例。

常见的社交软件功能在本研究中的应用说明如表 1 所示。

①动态: Facebook 会根据装置的浏览偏好、所在位置、好友信息、个人信息等已知信息事先产生好要传送给装置的资料;

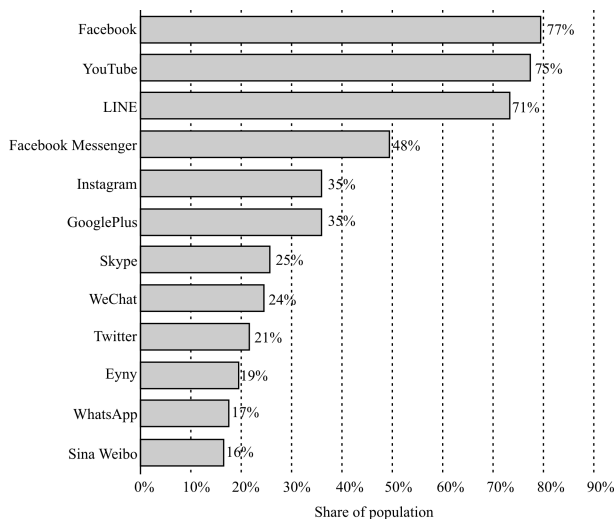


图 1 2016 年台湾地区常用的社交软件统计图

(资料来源: <https://www.statista.com/statistics/295611/taiwan-social-network-penetration/>。)

②信息：Facebook 可以得知该账号尚未读取的信息；

③绑定手机号码：根据装置的信号定位位置，让基地台知道邻近的装置。

表 1 常见社交软件功能

	Line	Instagram	Facebook	Google+
广告/动态	✓/✓	✓/✓	✓/✓	✓/×
讯息/聊天室群组	✓/✓	✓/×	✓/✓	✓/×
绑定手机号码	✓	不一定	不一定	不一定

资料来源：本研究整理。

## 1.2 研究动机

内容传递网络（content delivery networks, CDN）<sup>[4]</sup>是指通过网络进行相连的网络系统，使用最邻近装置的服务器，传送信息给装置，进而提高系统效能，增加可扩展性与低成本的内容，如图 2 所示。

在 5G 底下的内容传递网络（CDN）（如图 3 所示）运作模式包括：

①将资料传送到网络中的某一个区域主机，作为专属运作的服务器；容量与处理速度快，但是和装置的传输距离较远。

②将资料暂存于 Cloud RAN（cloud of radio access network，云无线接入网）。

③将资料快取于远端无线电端（RRH），在装置送出请求时回传结果；距离装置的装

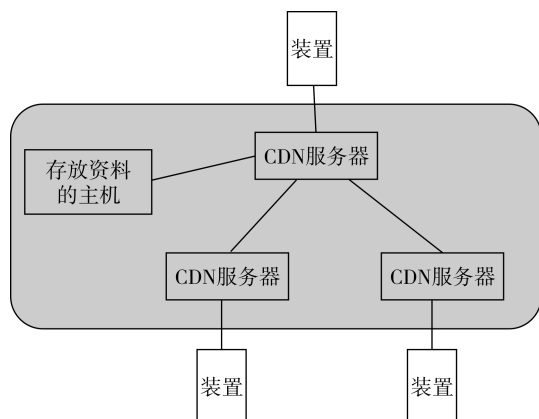


图 2 内容传递网络示意图

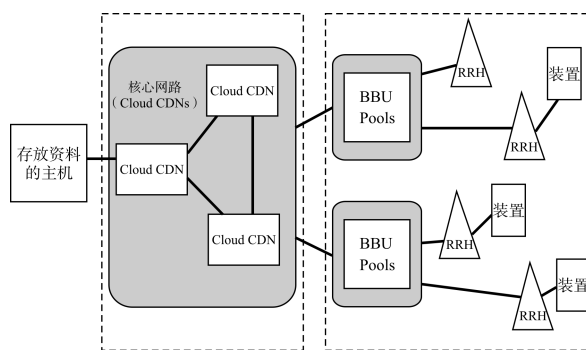


图 3 5G 底下内容传递网络示意图

置较近，但是容量小，所能存取符合装置需要的资料有限。

本研究中将资料先快取于远端无线电端（RRH），考虑在协同多点传输（coordinated multipoint, CoMP）的运作之下，把不同的资料快取（caching）于不同的远端无线电端（RRH）之中<sup>[7]</sup>，使用雾运算（fog computing）增加资料储存容量，再根据装置请求直接传送给装置，进而提升服务效率以最小化回应时间增加请求资料命中率（hit rate）。

## 1.3 研究目的

根据以上的研究背景与动机，本研究目的为：

装置请求的资料量往往大于基地台内的快取空间，在空间有限的情况下，通过已有的信息对装置可能请求的资料进行决策，判

断哪些资料优先存放于基地台快取空间中；在装置连线的基地台预先存放装置可能会存取的资料，让装置不用到远端主机存取资料，当装置发出请求后直接从基地台传输到装置，提高命中率，减少资料传输时间延迟，有效利用基地台内可存放快取资料的空间。

## 2 文献探讨

### 2.1 5G

5G (5th generation mobile networks 或 5th generation wireless systems) 指的是第五代移动通讯技术<sup>[5]</sup>，也是 4G 之后的延伸，但是目前对于 5G 的全球标准尚未确定，因此各家厂商对于 5G 的定义也略有不同。下一代行动网络联盟 (next generation mobile networks alliance) 定义 5G 为<sup>[6]</sup>：

①以 10Mbps 的数据传输速率支持数万户用户，同时提供 1Gbps 的数据传输速率。

②支持数十万的大量连线以用于支持大规模传感器网络的部署。

③频谱效率、覆盖率和 4G 相比有显著增强，延迟率也比 LTE (通用移动通信技术的长期演进) 低。

未来 5G 技术最有可能使用的频谱为

28GHz 及 60GHz，属极高频 (EHF)，比现在一般通信使用的频谱高出许多。虽然 5G 能提供极快的传输速度，但是高频带来的问题是传输距离比较短，穿墙能力有限，因此需要通过兴建大量密集的基地台覆盖以解决这个问题，在大量部署基地台的情况下若可以将资料事先放入较接近装置的基地台快取则可以减少传输时间。

### 2.2 快取

快取 (cache) 这项技术在网络领域中已经存在很久了，这样的技术是当装置浏览过该网页后，将网页内容暂存在装置中，下次存取相同网页的时候可以直接从本机端读取资料，不必再到远端的主机请求传回资料，节省时间和频宽。

由于现在无线网络流量爆炸性成长，将快取技术应用在无线网络中可以减少回程负载以及传输时间。

### 2.3 PHY-caching

PHY-caching (physical layer-caching)<sup>[8,9]</sup> 是一种在无线网络中可以利用基地台的快取减轻干扰以及提高无线电接口中的自由度 (degrees of freedom, DoF) 的快取技术 (如图 4 所示)。

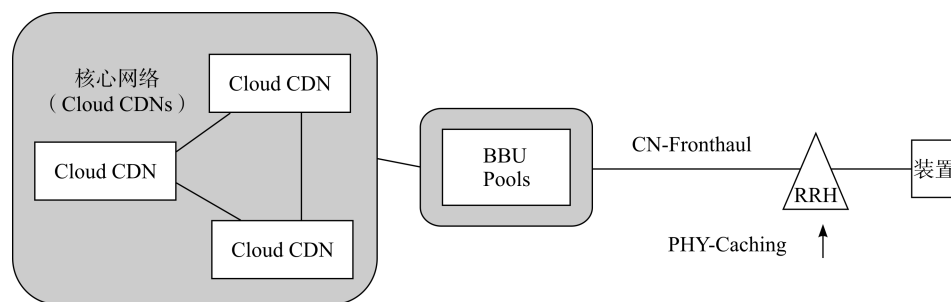


图 4 PHY-caching 示意图

如果同时有很多装置一起存取邻近基地台的内容，它可以减少基地台内的动态信息，促使它们可以用协同多点传输 (coordinated multipoint, CoMP) 服务增益而不消耗 CN-Fronthaul (gateway 和基地台之间) 的频宽，如此一来，可以有效地转变成多输入多输出 (multiple-input multiple-output, MIMO) 广播拓扑。

## 2.4 雾运算

“雾运算”这个概念由思科 (Cisco) 首创<sup>[10,11]</sup>，是一种水平的点到点架构，能够大量运算、储存资源和服务直接传送给装置或邻近的装置，类似云端运算，但不同于云端运算采用集中式处理资料。雾运算采用分散式的运算方式，比云端运算服务范围更大，在一个或多个装置或者是邻近的装置储存资

料，而不是全部将资料放在远端的主机，更接近使用者<sup>[12]</sup>，任何装置若具备连网、运算以及储存功能，就能够成为一个雾运算的节点。(如图 5 所示) 雾运算可以使资料传输降低延迟时间，传输范围也更大，具有更强的行动性<sup>[13]</sup>，在本研究中运用雾运算在特定基地台服务范围储存快取资料，弥补在单一基地台内空间不足的缺点。

## 2.5 协同多点传输

CoMP<sup>[14,15]</sup>是一种传输和接收的技术，利用多个天线传输和接收从基地台传送的资料，增强收到的信号质量和减少空间的干扰。(如图 6 所示) CoMP 在本研究中的运用中将会和前面提到的雾运算结合，在基地台找到资料后通过 CoMP 传回给装置。

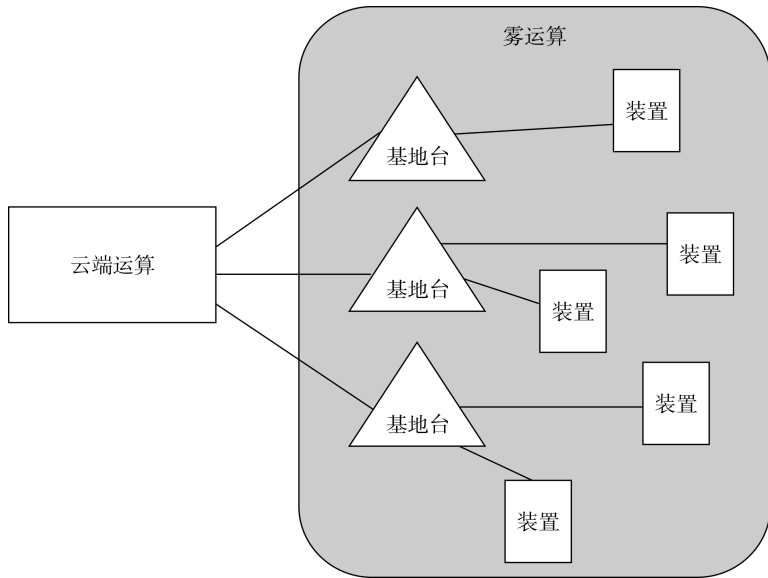


图 5 雾运算示意图

## 2.6 代理服务器、快取、内容传递网络的说明比较

代理服务器设定有一个较大的缓冲区，当有网际网络资料通过时代理服务器会将其

资料储存在缓冲区中，当其他使用者要存取相同资料时，代理服务器先检查自己的快取中有没有该资料，有的话直接会从缓冲区取出资料传给使用者，没有的话会帮装置去向远端取得装置所需要的资料。当装置在快取

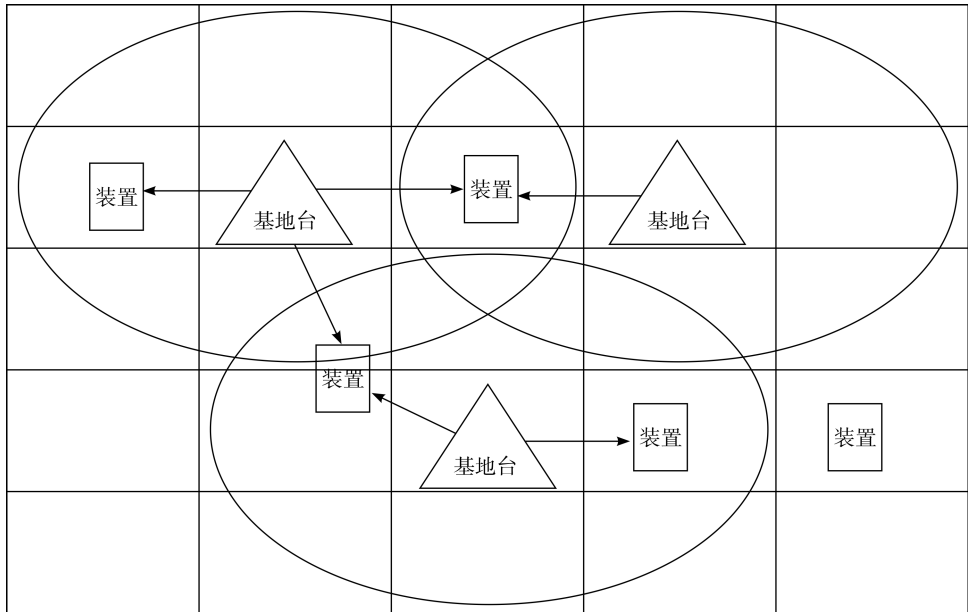


图 6 CoMP 示意图

中找不到资料时,代理服务器会代为向远端取得资料。

差异:代理服务器必须先存取过资料才会将存取过的内容存在代理服务器里。

①应用程序的快取:将常被存取的资料放在邻近的基地台待需要存取时直接存取不必再到远端找资料。

②传统的快取:将装置需要的资料先储存在本机端中,待装置下次要存取时就可以直接存取不必再到远端找资料。

差异:传统的快取一定要先存取过才能存在本机端中。

内容传递网络 (content delivery networks, CDN):使用最靠近装置的服务器,将资料暂存在该服务器中,待装置发出请求时再传送给装置。

差异:在提出要求前没有被存取过的资料也可以储存在该服务器中。

### 3 研究方法

#### 3.1 问题说明

基地台中只有一部分的空间可以存放快取资料,根据连线的装置数量可能会产生以下情况:

①请求的资料量大于基地台可储存的快取资料容量。

②请求的资料量等于基地台可储存的快取资料容量。

③请求的资料量小于基地台可储存的快取资料容量。

在装置资料请求量小于或等于基地台可储存的快取资料容量时可以将资料全部快取于基地台内,但是在资料请求量大于基地台可储存的资料容量时则必须做出取舍,选择可以摆放的资料存放,尽可能提高资料命中率。

本研究主要探讨在资料请求量大于基地台快取资料容量的情况,通过已知信息对要摆放在基地台内的快取资料进行选择,决定



哪些资料可以放进基地台的快取空间。

### 3.2 情境说明

对基地台、装置和资料的已知信息说明：

#### (1) 基地台

①已连线的装置：得知要摆放哪些可能被装置请求的资料。

②快取资料的容量：只能在基地台内存放几个资料。

#### (2) 装置

①连线到的基地台：得知快取资料被摆放的地方。

②存取资料频率：装置长期以来在  $\pi$  时段请求资料的记录，根据这些记录得知装置可能会请求资料的概率中。

③请求的资料：假设筛选要给装置的快取资料中，有资料在相同基准点上重复被多个装置存取的，则可以优先考虑摆放在基地台中。

#### (3) 资料

①资料类型：在相同的条件下假设资料大小都一样，站在使用者的角度思考使用者对于延迟的敏感度影音会大于文字，因此如果在同样的时间请求资料，优先权为“影音 > 图片 > 文字”。

②资料大小：可以得知假设资料在替换前未被存取的成本。

事件说明

#### (1) 初始状态

基地台得知已连线的装置，并将该装置可能请求的资料进行概率权重排序，预先将资料存放于基地台快取空间内。

#### (2) 资料替换

因应连线装置的资料量大于基地台的快取空间，在空间有限的情况下快取资料必须要定期替换。

#### (3) 装置请求资料被读取

当装置存取过基地台内属于它的快取资料后，该快取资料会先检查该资料是否还会被其他装置存取以便做决策，要留下或者是剔除。

#### (4) 新装置连线或装置断线

当新装置连线或装置断线时，会比对事件发生前后的基地台内的资料顺序做最佳决策，决定是否要替换基地台内资料顺序。

## 4 结语

针对 Facebook 的基地台资料部署方式为 Facebook 已事先产生好要传递给该装置的内容，基地台通过已知信息做出决策，将连线到该基地台的所有装置可能会存取的资料事先快取在基地台中，已知信息有辨别哪些装置连到基地台、装置所在地点、装置习惯存取资料的时间等，在基地台中事先放入 Facebook 已准备好装置可能会请求的资料，这个基地台部署资料方式虽然不能涵盖所有种类的应用程序，但是可以适用于跟 Facebook 相似的社交网络软件，例如 Line 或 Instagram。

## 参考文献

- [1] G Paschos, E Bastug, I Land, et al. Wireless caching: Technical misconceptions and business barriers [C] //IEEE Communications Magazine, 2016, 54 (8): 16—22.
- [2] E Bastug, M Bennis, M Debbah. Living on the edge: The role of proactive caching in 5G wireless networks [C] //IEEE Communications Magazine, 2014, 52 (8): 82—89.

- [3] T S Portal. Penetration of leading social networks in Taiwan as of 4th quarter 2016 [ R/OL ]. 2017. <https://www.statista.com/statistics/295611/taiwan-social-network-penetration/>.
- [4] J Tang, T Q Quek. The role of cloud computing in content-centric mobile networking [ C ] //IEEE Communications Magazine, 2016, 54 (8): 52—59.
- [5] J. G. Andrews, et al. What will 5G be? [ C ] //IEEE Journal on selected areas in communications, 2014, 32 (6): 1065—1082.
- [6] N Alliance. NGMN 5G White Paper V1.0 [ C ] //approved and delivered by the NGMN Board, 2015.
- [7] X Wang, M Chen, T Taleb, et al. Cache in the air: exploiting content caching and delivery techniques for 5G systems [ C ] //IEEE Communications Magazine, 2014, 52 (2): 131—139.
- [8] A Liu, V K Lau. How much cache is needed to achieve linear capacity scaling in backhaul-limited dense wireless networks? [ C ] //IEEE/ACM Transactions on Networking (TON), 2017, 25 (1): 179—188.
- [9] W Han, A Liu, V K Lau. PHY-caching in 5G wireless networks: Design and analysis [ C ] //IEEE Communications Magazine, 2016, 54 (8): 30—36.
- [10] Y J Ku, et al. 5G Radio Access Network Design with the Fog Paradigm: Confluence of Communications and Computing [ C ] //IEEE Communications Magazine, 2017, 55 (4): 46—52.
- [11] M Chiang, S Ha, I Chih-Lin, et al. Clarifying fog computing and networking: 10 questions and answers [ C ] // IEEE Communications Magazine, 2017, 55 (4): 18—20.
- [12] J Pan, J McElhannon. Future edge cloud and edge computing for internet of things applications [ C ] //IEEE Internet of Things Journal, 2018, 5 (1): 439—449.
- [13] H Liu, F Eldarrat, H Alqahtani, et al. Mobile edge cloud system: Architectures, challenges, and approaches [ C ] //IEEE Systems Journal, 2017 (99): 1—14.
- [14] R Irmer, et al. Coordinated multipoint: Concepts, performance, and field trial results [ C ] //IEEE Communications Magazine, 2011, 49 (2): 102—111.
- [15] D Lee, et al. Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges [ C ] //IEEE Communications Magazine, 2012, 50 (2): 148—155.