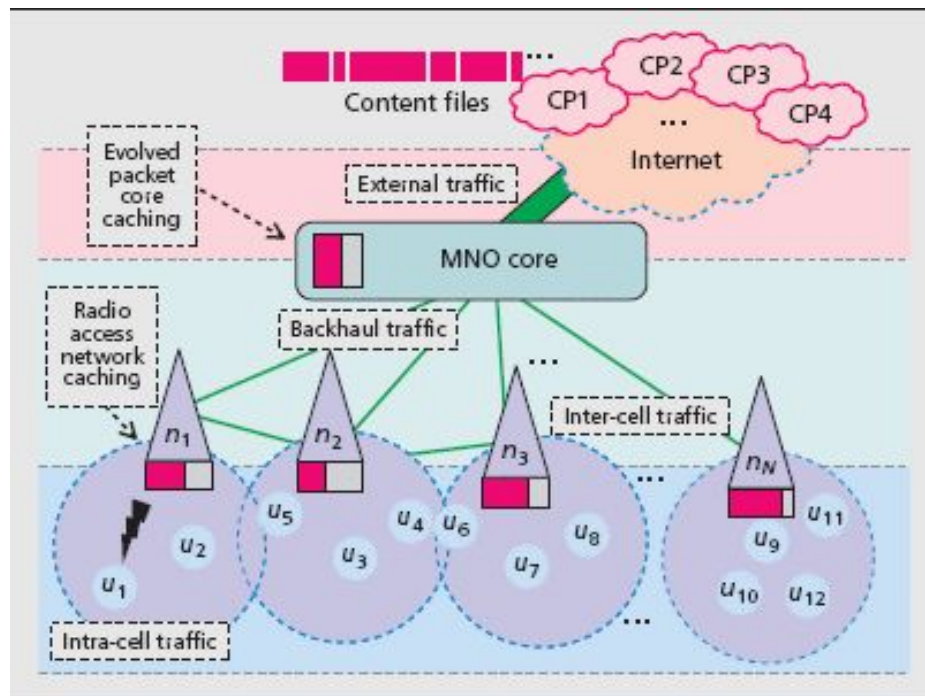


## Existing model



*Sight modification:* To implement **Redundancy Elimination** in caching, as said in “Cache in the air: exploiting content caching and delivery techniques for 5G systems” page 4. This would help removing duplicate transmissions.

It involves deploying 2 *middle boxes*, one on the link between between Internet (ISP) and MNO core(Main base station), and the other on the links between MBC and RAN (small base station).

I have some metrics based on which I can calculate the efficiency of the design

- Satisfied requests : “*Satisfied*” can be defined as the requested content being delivered with less than a threshold delay time.
- Requests routed to MBC and ISP: should be minimized.
- Hit Rate
- Delay in retrieving user requested content

These metrics depend on some parameters:

- Cache storage capacity: assumed same for all RAN, which is less than that of MBC.
- Content Transmission capacity: this makes it more realistic, as during peak traffic, RANs cannot answer all requests routed to them, even if they have the relevant content.
- User demand heterogeneity.

The caching and routing policy should achieve these objectives:

- Maximizing the *satisfied* user requests
- Decreasing the delay time in content transmission.
- Decreasing the inter-ISP traffic ( traffic between internet and various MBCs).
- Decreasing the intra-ISP traffic (traffic between RANs and their common MBC)

The last two demands are often contradicting, so an optimal tradeoff is required between them.

To begin with, this caching and routing policy could be a simple greedy one, which may not satisfy all the objectives.