A General Performance Evaluation Framework for Network Selection Strategies in 3G-WLAN Interworking Networks

Hao Wang¹, Dave Laurenson¹, and Jane Hillston²

¹Institute for Digital Communications
²Laboratory for Foundations of Computer Science
The University of Edinburgh
Outline

- 3G-WLAN Interworking Networks and Network Selection Strategies
- Models of Network Selection Strategies
- Derivation of Network Blocking Probabilities and Handover Rates
- Evaluation Results
- Conclusions
Heterogeneous wireless networks

- Users are able to use a wide range of wireless networks, often with multiple networks available at the same time.
Heterogeneous wireless networks

- Heterogeneous wireless networks have complementary characteristics such as data rate and coverage, e.g.

<table>
<thead>
<tr>
<th></th>
<th>Coverage Area</th>
<th>Data Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>3G</td>
<td>~ 1 – 2 km</td>
<td>2 Mbps (3G)</td>
</tr>
<tr>
<td>WLAN</td>
<td>~ 100 – 200 m</td>
<td>54 Mbps (802.11a)</td>
</tr>
<tr>
<td>Bluetooth</td>
<td>~ 10m</td>
<td>24 Mbps (version 3.0)</td>
</tr>
</tbody>
</table>

Therefore, it is envisioned that next-generation wireless communications will focus on the integration of these heterogeneous networks.
3G-WLAN interworking architecture

- It is becoming necessary to integrate wireless LANs (WLANs) and 3G cellular networks, to form 3G-WLAN interworking networks.
In heterogeneous wireless networks, a mobile node may perform handovers during its communications:

- horizontal handover (HHO): a mobile node moves across cells that use the same type of access technology.
- vertical handover (VHO): the movement between different types of wireless networks.
Handover decision of HHO and VHO

- Before a mobile node performs either handover it must:
  - collect information to confirm the need for a handover, and
  - decide whether to perform the handover.

- For a HHO, the handover criterion is usually just the signal strength received by the mobile node.

- For a VHO, various handover criteria can be taken into account when making a handover decision e.g.:
  - **cost of service**: cost is a major consideration, and could be sometimes be the decisive factor.
  - **network conditions**: network-related parameters such as bandwidth and network latency.
  - **mobile node conditions**: the node’s dynamic attributes such as mobility pattern, account balance and power consumption.
  - **user preference**: a user may have preference for one type of network over another.
To facilitate the above evaluation process, mathematical expressions are introduced: network selection strategies (NSSs).

A number of NSSs have been proposed and they are generally based on multiple attribute decision making (MADM) theory.

A typical example is the simple additive weighting (SAW) strategy:

Each network is associated with a point, which is calculated as the weighted sum of the handover-related attribute values.

\[
P_i = \sum_{j=1}^{N} w_j \times r_{ij}, \text{ where } r_{ij} = \frac{x_{ij}}{\sum_{i=1}^{M} x_{ij}}
\]

- there are \( N \) attributes
- weight of attribute \( j \)
- normalised value of attribute \( j \) of network \( i \), where \( M \) is the number of candidate networks.
- this is used to cancel the effect of the unit of different attributes
Framework structure

- Captures movement characteristics in 3G-WLAN environment
- Represents features of multimedia services
- Controls network selection behaviour of a mobile node
- The generality of the framework is achieved by having two interfaces
- Determines network selection probabilities
- Represents how network resources are consumed by mobile nodes

- Network resource consumption model
- Network selection strategy
Traffic model

The traffic model of a mobile node is modelled in the session model, which includes two parameters: session arrival rate and session duration.

Field data suggests that the statistical session duration of multi-type-services has a coefficient of variation (CoV) larger than one.

To capture this feature, we use the hyper-exponential distribution (HED) to model the session duration. A two-phase HED is used in this work, where one phase represents non-real time (NRT) sessions and the other represents real time (RT) sessions.
Traffic model

- As for session arrival rate, the general consensus that the session arrival is a Poisson process is followed.

- The traffic model is constructed as a combination of two ON-OFF sources:
In 3G-WLAN interworking networks, a 3G cellular cell is generally overlaid with one or more WLAN cells.

The mobility model characterises a node’s residence time in:
- both the whole 3G-WLAN compound cell
- and different radio access technologies.

Thus a Coxian distribution is used as the mobility model:
- a K-phase Coxian structure is composed of a series of K exponentially distributed states and an absorbing state.

- it can approximate any probability distribution arbitrarily closely
Mobility model

- **even phases**: 3G-WLAN
  - transitions back to phase 1: movements out of a compound cell and entering another one

- **odd phases**: 3G only
  - transitions between neighbouring phases: movements between different RAT areas

Two assumptions are made:
- WLAN cells do not overlap with each other;
  - HHO between WLAN cells is not considered
- WLAN cells that overlap with adjacent cellular cells belong to all the cellular cells;
  - the start point of the track of the mobile node in a 3G-WLAN compound cell is always the 3G area
The above mobility model can capture various traces of the mobile node in 3G-WLAN interworking networks.

- **Trace 1**: phase 1 > phase 1
- **Trace 2**: phase 1 > phase 2 > phase 1
- **Trace 3**: phase 1 > phase 2 > phase 3 > phase 1
- **Trace 4**: phase 1 > phase 2 > phase 3 > phase 1
In the PEPA model for NSSs, a mobile node
- can generate different types of sessions, and these sessions are submitted to different networks according to NSSs (parameters $P_C$ and $P_W$ are used in the definitions of PEPA models);

- can perform different types of handovers according to the NSSs;

- is aware of network blocking for both new and handover sessions in 3G and WLAN networks (parameters $P_{B_C}$ and $P_{B_W}$ are used in the definitions of PEPA models);

- is aware of the different data rates that are provided by different RATs; (NRT sessions (e.g. file downloading) usually need less time using WLAN RAT than using 3G RAT)
In this work, a system state of a PEPA model is denoted as:

- $A$, the network the mobile node is connected to
- $B$, the type of the session the mobile node is engaged in
- $k$, the mobile node’s phase of its mobility model

Three performance measures are investigated:

- average throughput
- handover rate
- network blocking probability
the percentage of time the mobile node spends using different RATs for different types of sessions.

\[ T_{C,NRT} = \sum_{i=1}^{N} \pi(s_{i}^{C,NRT}), \quad T_{C,RT} = \sum_{i=1}^{N} \pi(s_{i}^{C,RT}), \]

\[ T_{W,NRT} = \sum_{i=1}^{N/2} \pi(s_{2i}^{W,NRT}), \quad T_{W,RT} = \sum_{i=1}^{N/2} \pi(s_{2i}^{W,RT}), \]

then, calculate the total engaged time of the mobile node:

\[ T_{Engaged} = T_{C,NRT} + T_{C,RT} + T_{W,NRT} + T_{W,RT} \]
then, the average throughput is defined as a weighted sum:

\[
THP = D_{NRT}^C \frac{TC, NRT}{T_{Engaged}} + D_{RT}^C \frac{TC, RT}{T_{Engaged}}
+ D_{NRT}^W \frac{TW, NRT}{T_{Engaged}} + D_{RT}^W \frac{TW, RT}{T_{Engaged}},
\]

the data rates that can be achieved using different RATs for different sessions
Handover rate is defined as the throughput of corresponding activities. States that can perform the corresponding handover have activity rates.

Handover rates:

- \( r_{C-C}^{\text{inter}} \):
  \[
  r_{C-C}^{\text{inter}} = \sum_{i=1}^{N/2} \left( b_{2i-1} * \nu_{2i-1} * \pi(s_{2i-1}^{C}) + b_{2i} * \nu_{2i} * \pi(s_{2i}^{C}) \right)
  \]

- \( r_{W-C}^{\text{inter}} \):
  \[
  r_{W-C}^{\text{inter}} = \sum_{i=1}^{N/2} b_{2i} * \nu_{2i} * \pi(s_{2i}^{W})
  \]

- \( r_{C-W}^{\text{intra}} \):
  \[
  r_{C-W}^{\text{intra}} = \sum_{i=1}^{N/2} P_{W} * a_{2i-1} * \nu_{2i-1} * \pi(s_{2i-1}^{C})
  \]

- \( r_{W-C}^{\text{intra}} \):
  \[
  r_{W-C}^{\text{intra}} = \sum_{i=1}^{N/2} a_{2i} * \nu_{2i} * \pi(s_{2i}^{W})
  \]
Network blocking probability

- Like network selection probabilities, these network blocking probabilities can be used as input parameters.

- In this work, they are derived from a 2D-CTMC that models the resource consumption of a 3G-WLAN compound cell.

- The state of the 2D-CTMC is denoted by two integers \((c,w)\), where \(c\) and \(w\) represent the numbers of engaged users in 3G and WLAN networks respectively;
There are five types of events that can change the state of the 2D-CTMC:

- New session requests are generated in 3G and WLAN networks.
- Sessions are finished and resources are released.
- Sessions are externally handed over into 3G and WLAN.
- Sessions are internally handed over between 3G and WLAN.
- Sessions are externally handed over out of 3G and WLAN.
Network blocking probability

- note that the definition of the 2D-CTMC uses handover rates as parameters.
The blocking probabilities of 3G and WLAN networks are then calculated as:

\[
P_B^C = \sum_{0 \leq w \leq N_W} \sum_{c=N_C} p(c, w), \quad P_B^W = \sum_{0 \leq c \leq N_C} \sum_{w=N_W} p(c, w),
\]
An implicit problem

- As presented above, the derivation of network blocking probabilities from the 2D-CTMC model requires handover rates as input parameters.

- On the other hand, to derive handover rates from the PEPA models, network blocking probabilities are needed.

\[
F_{CTMC}(R_H) : \quad R_H \xrightarrow{f_{CTMC}} p(c, w) \xrightarrow{f_{PEPA}} \pi(s_k^{A,B}) \xrightarrow{f_{PEPA}} P_B, \quad P_B = [P_B^C, P_B^W]
\]

\[
R_H = [r_{intra}^{C-W}, r_{intra}^{W-C}, r_{inter}^{C-C}, r_{inter}^{W-C}]
\]
Convergence speed

- The convergence speed of the above iterative method is dependent on the parameter settings but very fast.
  - four types of NSSs have been investigated with 10 increasing session durations --- as the table shows in each case only a low number of iterations was needed.

<table>
<thead>
<tr>
<th>Model</th>
<th>Numbers of iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>[2, 2, 3, 4, 5, 7, 9, 11, 11, 13]</td>
</tr>
<tr>
<td>RRSS</td>
<td>[2, 2, 3, 4, 5, 7, 8, 11, 12, 13]</td>
</tr>
<tr>
<td>WLAN-first</td>
<td>[2, 2, 3, 4, 5, 6, 8, 10, 12, 13]</td>
</tr>
<tr>
<td>Service-based</td>
<td>[2, 2, 3, 4, 5, 6, 8, 10, 12, 13]</td>
</tr>
</tbody>
</table>

- moreover, the results of the method are NOT dependent on the initial values of network blocking probabilities
Evaluation Results
Four types of NSSs

- **Random**: the mobile node selects 3G and WLAN with equal probabilities, i.e., 0.5;

- **Relative received signal strength (RRSS)**: the mobile node selects the network with the strongest signal strength;

- **WLAN-first**: the mobile node always choose WLAN when it is available, because of its high bandwidth, small delay and cheap cost;

- **Service-based**: the mobile node selects 3G for RT sessions (for less handovers) and WLAN for NRT sessions (for high data rate);
Parameter settings

- Network selection probabilities of different NSSs are:

<table>
<thead>
<tr>
<th>Network Selection</th>
<th>$P_C$</th>
<th>$P_W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>RRSS</td>
<td>0.4</td>
<td>0.6 [10]</td>
</tr>
<tr>
<td>WLAN-first</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Service-based</td>
<td>$P_C = 1$ (for RT session)</td>
<td>$P_W = 1$ (for NRT session)</td>
</tr>
</tbody>
</table>
Controlled parameters

- Effect of two mobility patterns
  - mobility pattern 1 ($t_{3G-WLAN}=474$, $P_{NRT}=P_{RT}=0.5$)
  - mobility pattern 2 ($t_{3G-WLAN}=1200$, $P_{NRT}=P_{RT}=0.5$)
mobility pattern 1 \( (t_{3G-WLAN}=474, P_{NRT}=P_{RT}=0.5) \)

\[
THP = D_{NRT}^C \frac{T_{C,NRT}}{T_{Engaged}} + D_{RT}^C \frac{T_{C,RT}}{T_{Engaged}} + D_{NRT}^W \frac{T_{W,NRT}}{T_{Engaged}} + D_{RT}^W \frac{T_{W,RT}}{T_{Engaged}},
\]

- A larger WLAN selection probability also results in a higher throughput.
- For longer NRT sessions, SB can make better use of WLAN.
mobility pattern 2 ($t_{3G-WLAN}=1200$, $P_{NRT}=P_{RT}=0.5$)

- a longer sojourn time in 3G-WLAN area results in a higher throughput
mobility pattern 1 ($t_{3G-WLAN}=474$, $P_{NRT}=P_{RT}=0.5$)

It depends on how frequently 3G is chosen, and how long network resources are engaged.
mobility pattern 1 \((t_{3G-WLAN}=474, \ P_{NRT}=P_{RT}=0.5)\) and

mobility pattern 2 \((t_{3G-WLAN}=1200, \ P_{NRT}=P_{RT}=0.5)\)

A longer stay in 3G-WLAN area results in a higher 3G network blocking probability.
mobility pattern 1 ($t_{3G-WLAN}=474$, $P_{NRT}=P_{RT}=0.5$)

- again it depends on the traffic load in WLAN, and the difference is more obvious
- NRT sessions engage WLAN resources shorter than RT sessions, because their awareness of data rate
mobility pattern 1 (\(t_{3G-WLAN} = 474\), \(P_{NRT} = P_{RT} = 0.5\)) and mobility pattern 2 (\(t_{3G-WLAN} = 1200\), \(P_{NRT} = P_{RT} = 0.5\)).

- A longer stay in 3G-WLAN area results in a lower WLAN network blocking probability.
mobility pattern 1 ($t_{3G-WLAN}=474$, $P_{NRT}=P_{RT}=0.5$)

- it depends on the probability of being connected to WLAN when moving out a 3G-WLAN compound cell
- and also the WLAN selection probability
mobility pattern 2 ($t_{3G-WLAN}=1200$, $P_{NRT}=P_{RT}=0.5$)

Again, a longer stay in 3G-WLAN area results in lower handover rates.
traffic pattern 1 \( (t_{3G-WLAN} = 1200, P_{NRT} = 0.3 \, P_{RT} = 0.7) \)

- A larger WLAN selection probability also results in a higher throughput.
- For longer NRT sessions, SB can make better use of WLAN.
traffic pattern 2 ($t_{3G-WLAN}=1200$, $P_{NRT}=0.7$ $P_{RT}=0.3$)

A larger NRT session proportion results in a higher throughput difference between SB and WF gets smaller at larger NRT probability.
traffic pattern 1 ($t_{3G-WLAN}=1200$, $P_{NRT}=0.3$ $P_{RT}=0.7$)

- for SB, 70% traffic is injected in 3G network
- it depends on how frequently 3G is chosen, and how long network resources are engaged
traffic pattern 1 ($t_{3G\text{-}WLAN}=1200, P_{NRT}=0.3\ P_{RT}=0.7)$ and
traffic pattern 2 ($t_{3G\text{-}WLAN}=1200, P_{NRT}=0.7\ P_{RT}=0.3$)

- SB decrease by a larger extent as it is more sensitive on traffic type
- lower RT probability reduces 3G network blocking probability
traffic pattern 1 ($t_{3G-WLAN}=1200$, $P_{NRT}=0.3$ $P_{RT}=0.7$)

Again, it depends on the WLAN selection probability and WLAN resource engagement time.
traffic pattern 1 (\(t_{3G-WLAN}=1200, P_{NRT}=0.3 P_{RT}=0.7\)) and traffic pattern 2 (\(t_{3G-WLAN}=1200, P_{NRT}=0.7 P_{RT}=0.3\))

- for the others, larger NRT probability means more traffic is aware of high data rate of WLAN

- for SB, larger NRT probability increase traffic load in WLAN
traffic pattern 1 ($t_{3G-WLAN}=1200$, $P_{NRT}=0.3$ $P_{RT}=0.7$)
traffic pattern 2 \((t_{3G-WLAN}=1200, P_{NRT}=0.7 \ P_{RT}=0.3)\)

- For SB, larger NRT probability means WLAN is more frequently used, and thus higher handover rates.
Conclusions
In conclusion

- For deterministic strategies (service-based and WLAN-first):
  - easy to implement;
  - user knows which network is connected to;
  - their performance in terms of the investigated measures are usually the boundaries of the studies strategies;

- For non-deterministic strategies (RRSS and random):
  - not easy to implement
  - users experience uncertainty during handover;
  - they produce more balanced performance on the investigated measures;
Thank You!
mobility pattern 1 ($t_{3G-WLAN}=474$, $P_{NRT}=P_{RT}=0.5$)

- it depends on the probability of being connected to 3G when moving out a 3G-WLAN compound cell
mobility pattern 2 ($t_{3G-WLAN}=1200$, $P_{NRT}=P_{RT}=0.5$)

- A longer stay in 3G-WLAN area obviously reduces handover rates.
traffic pattern 1 ($t_{3G-WLAN}=1200$, $P_{NRT}=0.3$ $P_{RT}=0.7$)
traffic pattern 2 ($t_{3G-WLAN}=1200$, $P_{NRT}=0.7$ $P_{RT}=0.3$)

- A lower RT probability means less portion of time connected to 3G, thus lower handover rates.
- SB is more sensitive than the others, and is now almost the same as random.