Boston University
Computer Science department
Technical Report
BUCS-TR-2006-023

Georgios Smaragdakis
gsmaragd@cs.bu.edu

Notes on the Effect of Different Access Patterns on the Intensity of Mistreatment in Distributed Caching Groups

In this report, we extend our study of the intensity of mistreatment in distributed caching groups due to state interaction. In our earlier work [1], we analytically showed how this type of mistreatment may appear under homogeneous demand distributions. We provided a simple setting where mistreatment due to state interaction may occur. According to this setting, one or more "overactive" nodes generate disproportionately more requests than the other nodes. In this report, we extend our experimental evaluation of the intensity of mistreatment to which non-overactive nodes are subjected, when the demand distributions are not homogeneous.

The intensity of the mistreatment is captured by the normalized access cost and the normalized social access cost [1]. Under non-homogeneous demand distributions, the access pattern of the "overactive" node is the dominant factor in the increase of the intensity of the mistreatment in the caching group. Following the same methodology that was presented in [1] (with error tolerance=0.001), we measure the intensity of the mistreatment in the scenario where the nodes of the distributed caching group follow different access patterns.

The main contribution of this study is that the intensity of mistreatment is highly related to the access patterns of the nodes in the group. We use a plethora of different access pattern models for the "overactive" nodes, and we provide cases where the existence of non-homogeneous demand distributions increases the intensity of mistreatment.

We also take into consideration the rate imbalance of the "overactive" nodes as well the effect of the total size of the distributed caching group: we consider three different cases, (i) $C_{total} < N$, (ii) $C_{total} > N$ and (iii) $C_{total} \gg N$, where N is the total number of distinct objects.

We assume that the non-overactive nodes follow the same access patters and the "overactive" nodes deviate from these access patterns.

Remark 1: We can relax the above assumption, as the non-overactive nodes in the steady state they will follow the aggregated access pattern.

Remark 2: Recall that we assume that there is no flooding of the miss requests of the "overactive" nodes, thus such miss requests are satisfied by only one node. As a result the intensity of the mistreatment is not the worst possible (we have also captured this in our analysis).

1 Model 1

For non-overactive nodes we will maintain the popularity ranking of objects as it was $(o_1, o_2, \dots)$. For the overactive node, however, we will first reverse the popularity ranking, but will allow overlap with the popularity ranking of non-overactive nodes up to rank O. Let us call this rank offset. This model captures the locality of the most popular objects. In the following pages we plot the demand distributions of non-overactive and overactive nodes. If $O = 0$, the object demand distributions of non-overactive and overactive nodes are disjoint. On the other hand, if $O = N$, then the object demand distributions of non overactive nodes and the overactive nodes are identical.

Observation 1: If C_{total} is small, the intensity of the mistreatment is maximized when the object demand distributions (access patterns) are disjoint.

Observation 2: If C_{total} is high, the variance of the intensity of the mistreatment is very small for a fixed value of the offset.

Observation 3: The normalized Social cost follows the same trend as the normalized access cost of the group.

Please notice that the access cost of the overactive node (whose access pattern deviates from the access pattern of the group) is higher when the request rate is identical, and (slightly) decreases as the rate imbalance increases. This is commonly observed in most of the models we are studying in this report.

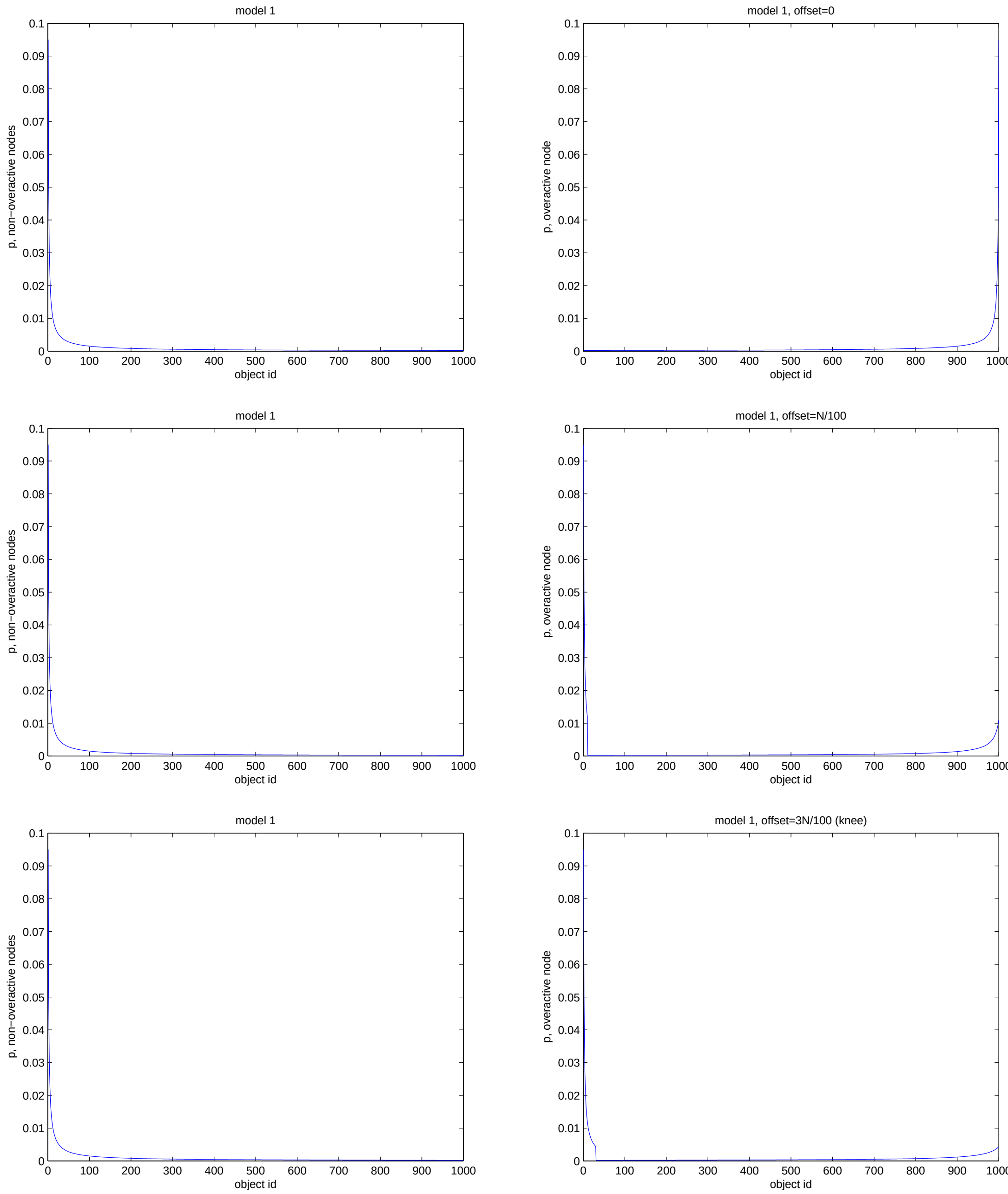


Figure 1: Model 1: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$\{0, N/100, 3N/100\}$.

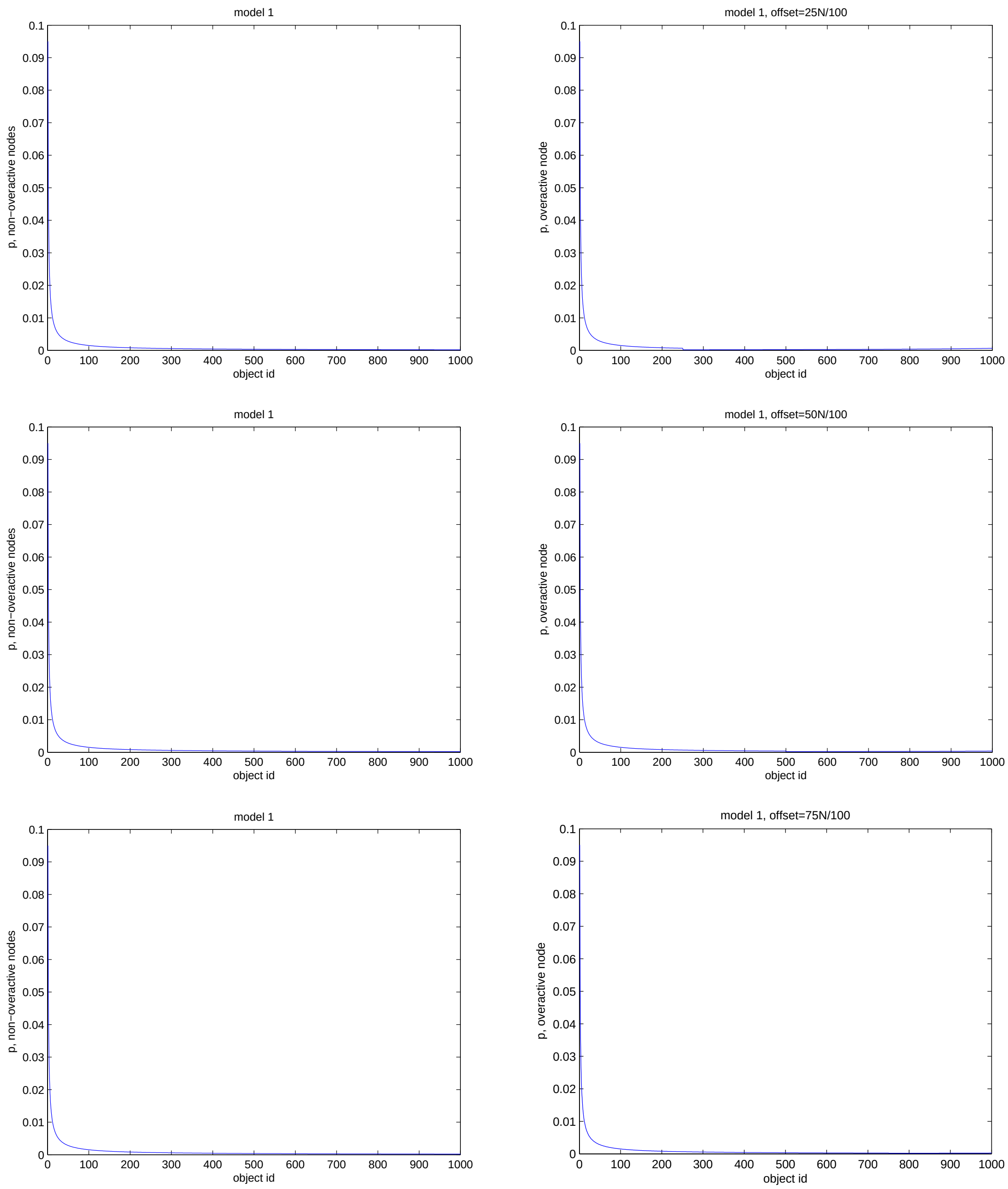


Figure 2: Model 1: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$\{25N/100, 50N/100, 75N/100\}$.

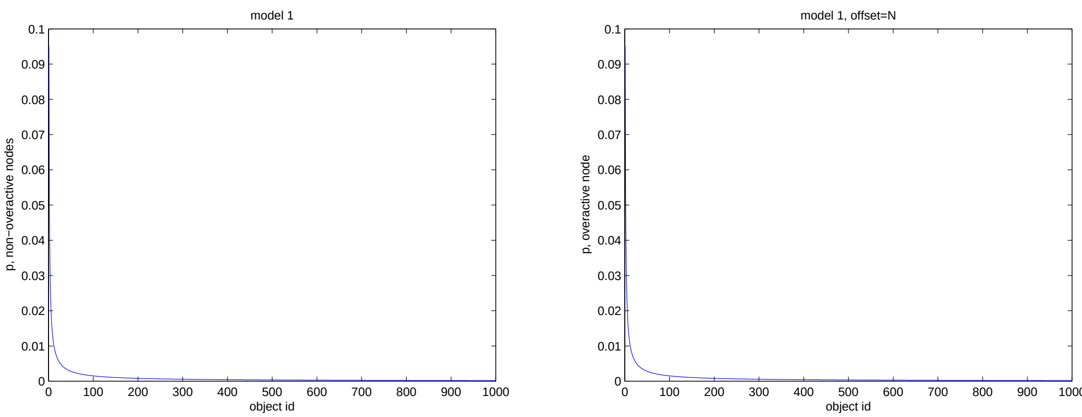
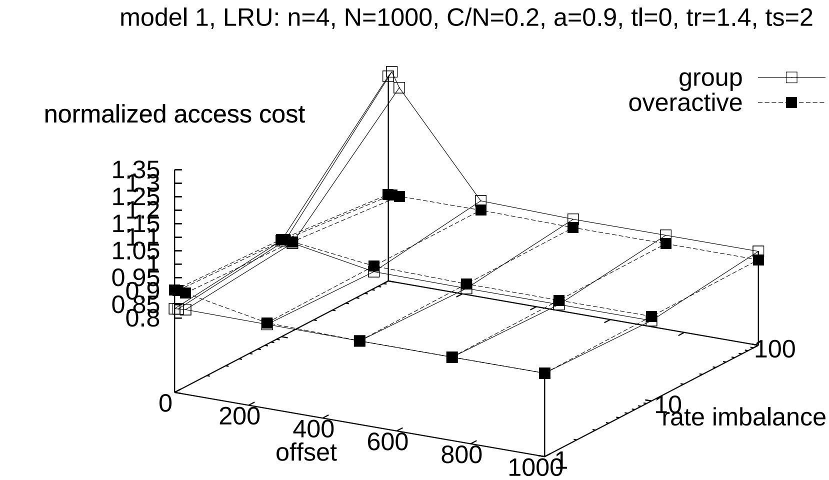


Figure 3: Model 1: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=N.



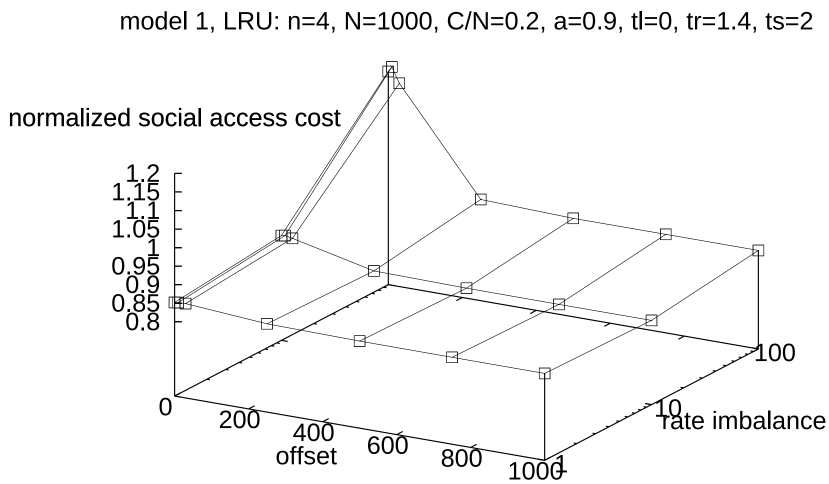
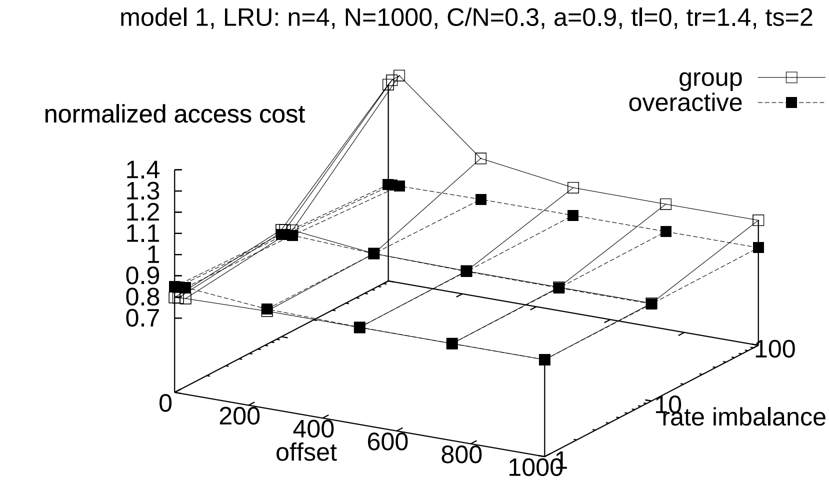


Figure 4: Model 1: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} < N$.



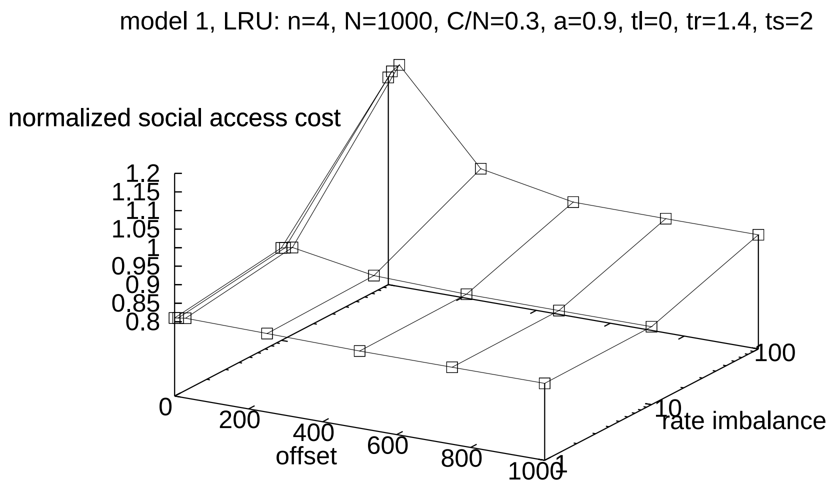
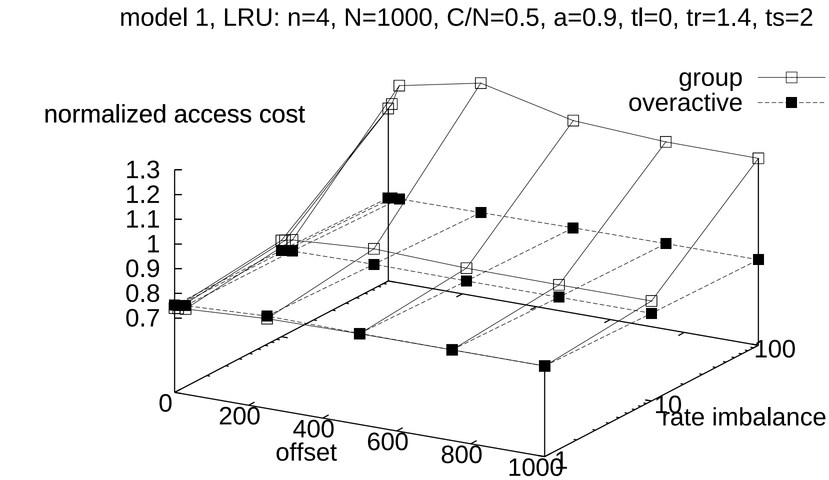


Figure 5: Model 1: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} > N$.



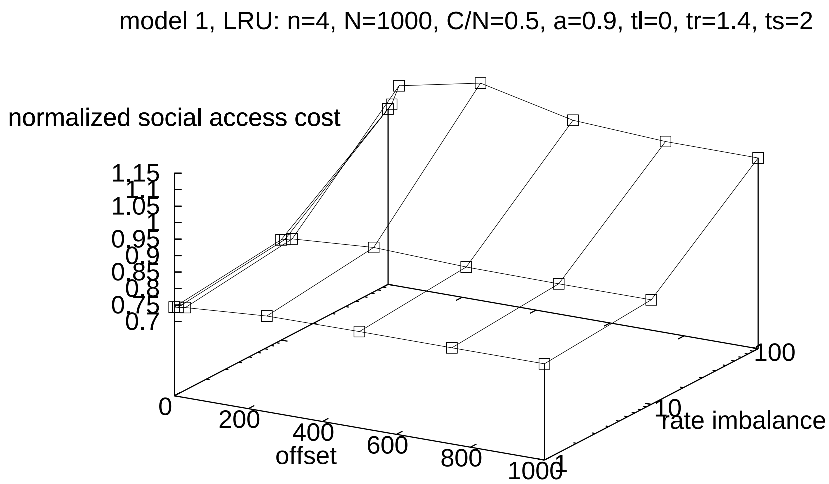


Figure 6: Model 1: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} \gg N$.

2 Model 2

For non-overactive nodes we will maintain the popularity ranking of objects as it was $(o_1, o_2, \dots)$. For the overactive node, however, we will shift the popularity ranking rightwards by an *offset* O, $0 \leq O \leq N$, therefore make object $o_{1+(O+i-1) \bmod N}$ be the ith most popular one. We assign request probabilities taken from the same generalized power-law profile with skewness $a = 0.9$ that is used for the non-overactive nodes.

We plot the demand distributions of the non-overactive and overactive nodes and the individual cost for the overactive and the non-overactive nodes as well as the social cost (the normalization conducted by dividing with the corresponding cost obtained when remote hits are not allowed to affect the local caching state). As it is obvious, mistreatments can occur even under non-homogeneous demand distributions.

Observations 1: When the C_{total} is small (compared to N), the concave profile with respect to O occurs as with high O the popularity ranking starts to look like the original one due to "wrapping" after N.

Observation 2: When the C_{total} is large (compared to N), the intensity of the mistreatment is maximized when the demand patterns of the non-overactive nodes and overactive nodes are identical. The profile of the normalized access cost is convex (taking the minimum for high values of O).

Observation 3: For all values of C_{total} we studied the normalized access cost and social access cost is lower than the one achieved when applying Model 1.

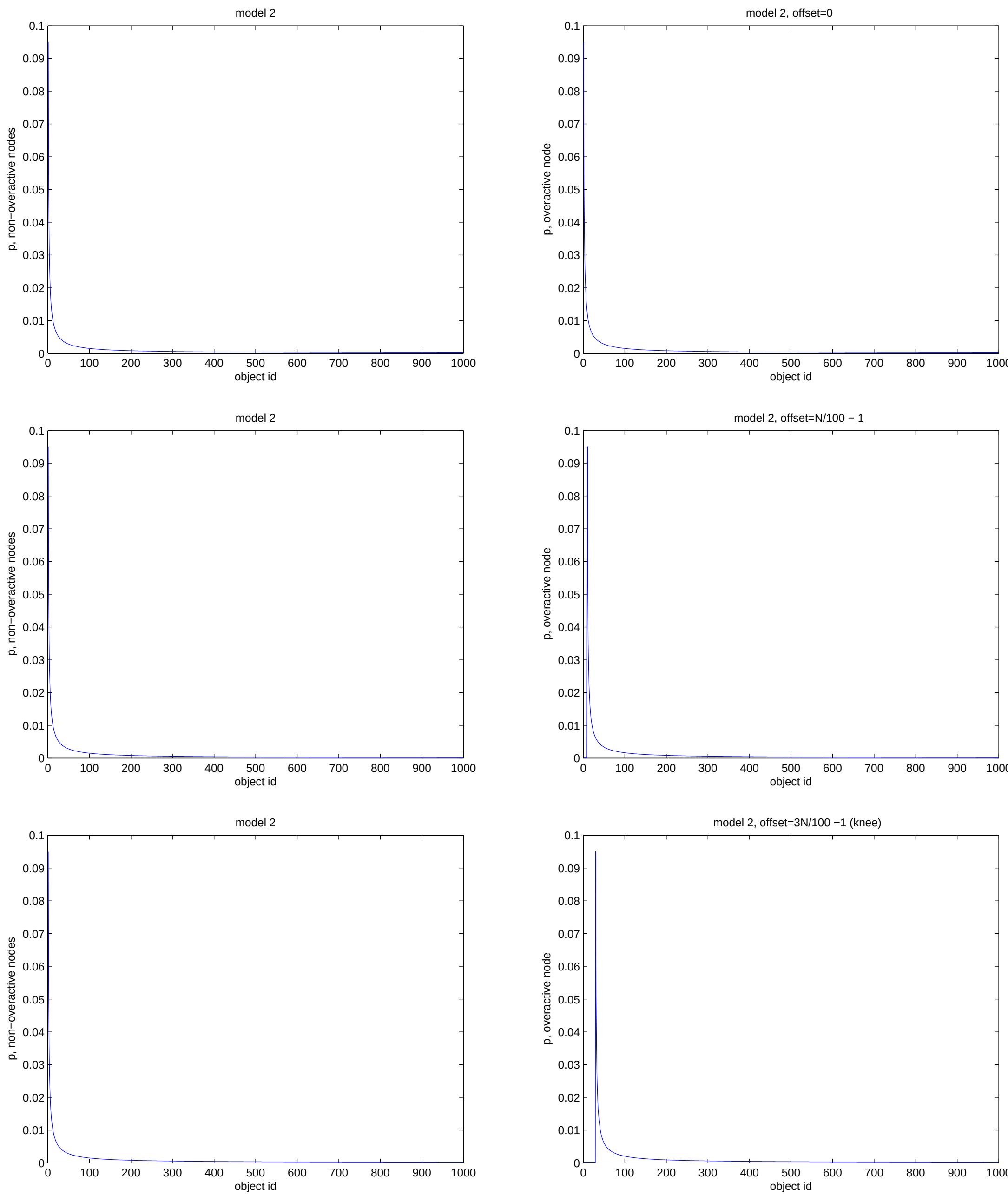


Figure 7: Model 2: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$\{0, N/100 - 1, 3N/100 - 1\}$.

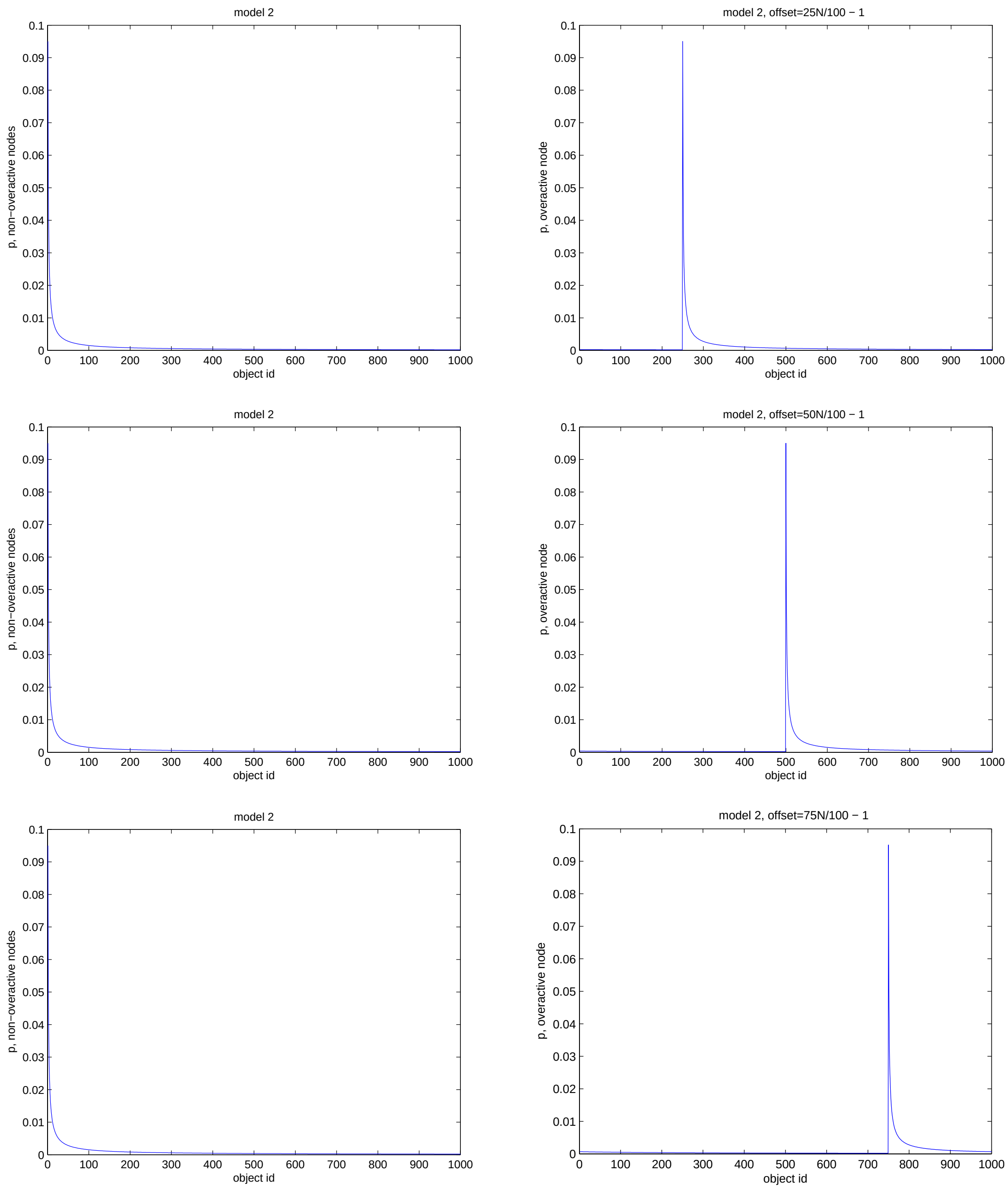


Figure 8: Model 2: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$\{25N/100 - 1, 50N/100 - 1, 75N/100 - 1\}$.

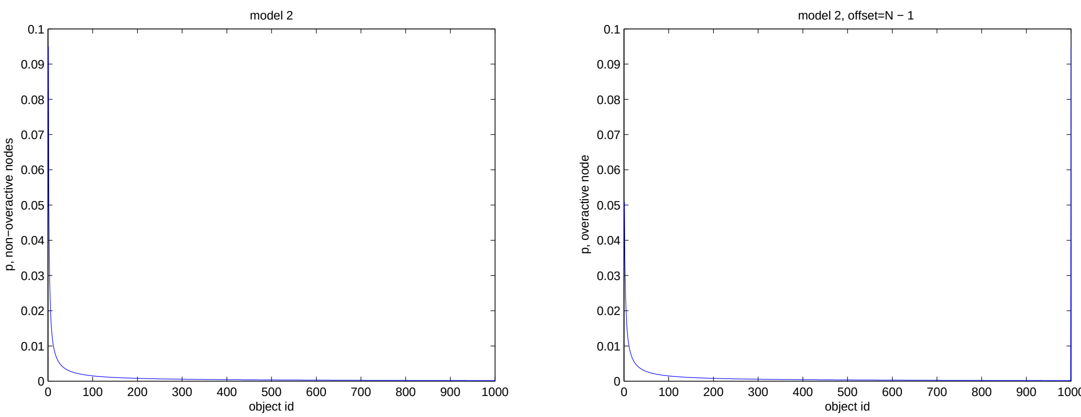


Figure 9: Model 2: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$N - 1$.

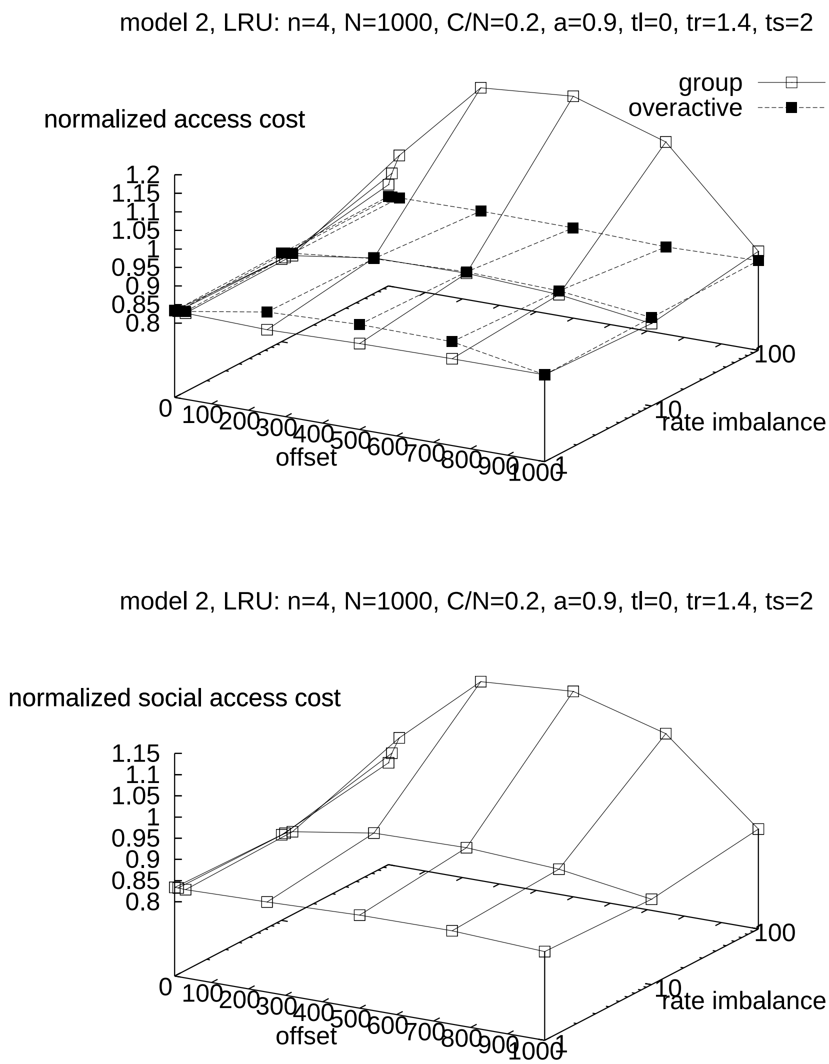
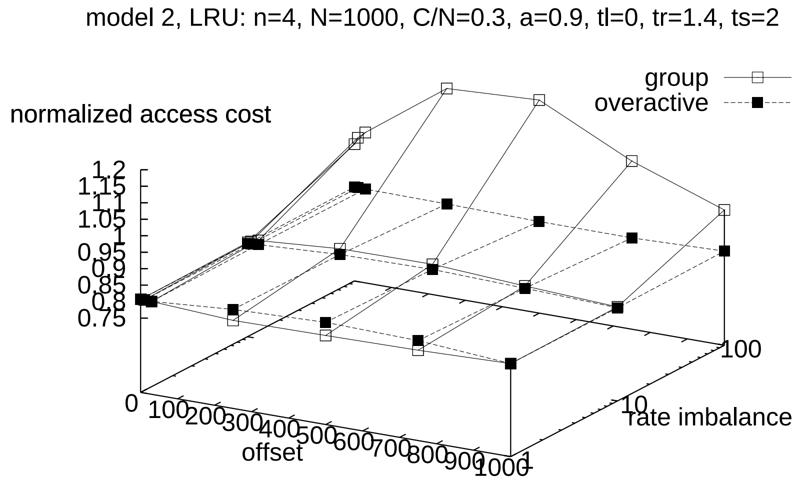


Figure 10: Model 2: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} < N$.



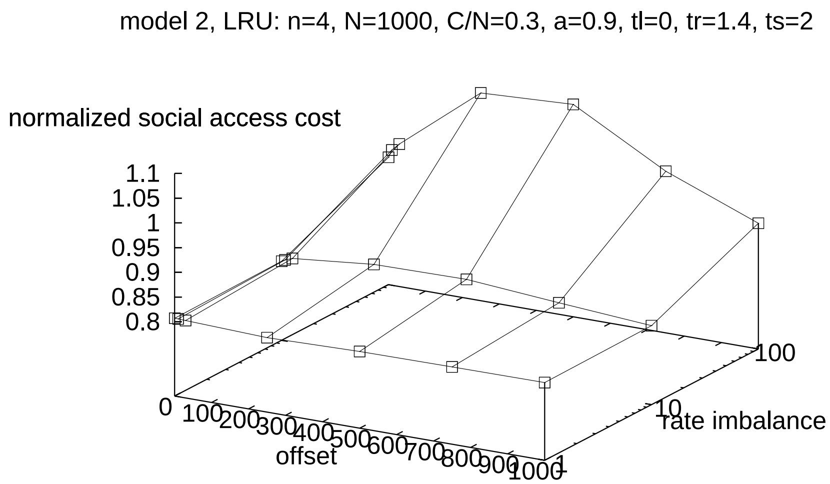
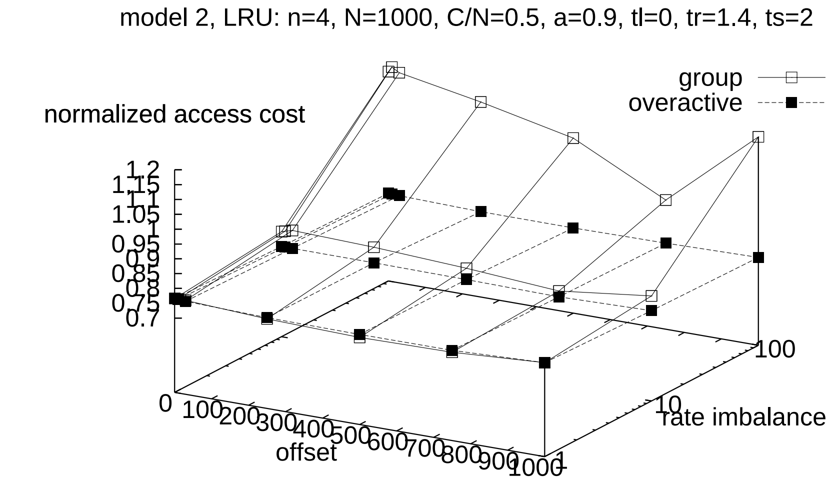


Figure 11: Model 2: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} > N$.



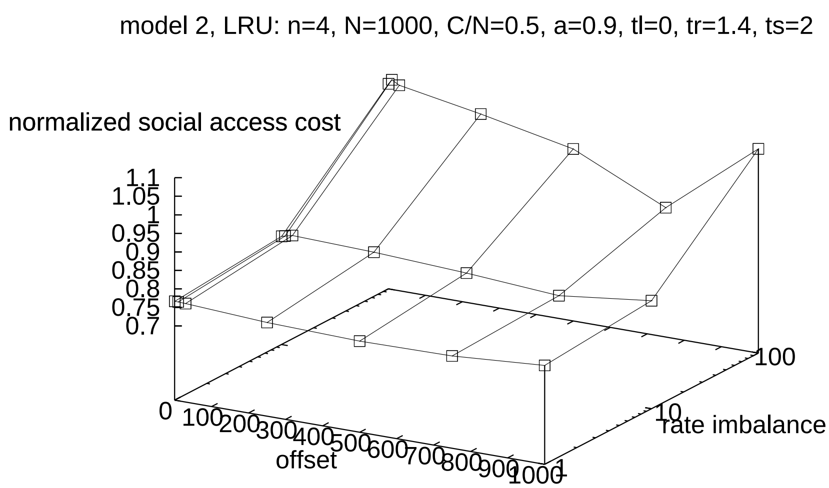


Figure 12: Model 2: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} \gg N$.

3 Model 3

Model 3 is similar to Model 2. The only difference is that when the overactive's popularity ranking of the objects wraps up, the rank is inversed, i.e. if the offset O is non-zero, then the loss popular object of the overactive node is the most popular node of the group.

Observation 1: When the C_{total} is small (compared to N), the intensity of the mistreatment increases monotonically with the offset.

Observation 2: When the C_{total} is large (compared to N), the variation of the intensity of mistreatment is very small for the same rate imbalance.

Observation 3: For all values of C_{total} we studied the normalized access cost and social access cost is lower than the one achieved when applying Model 1.

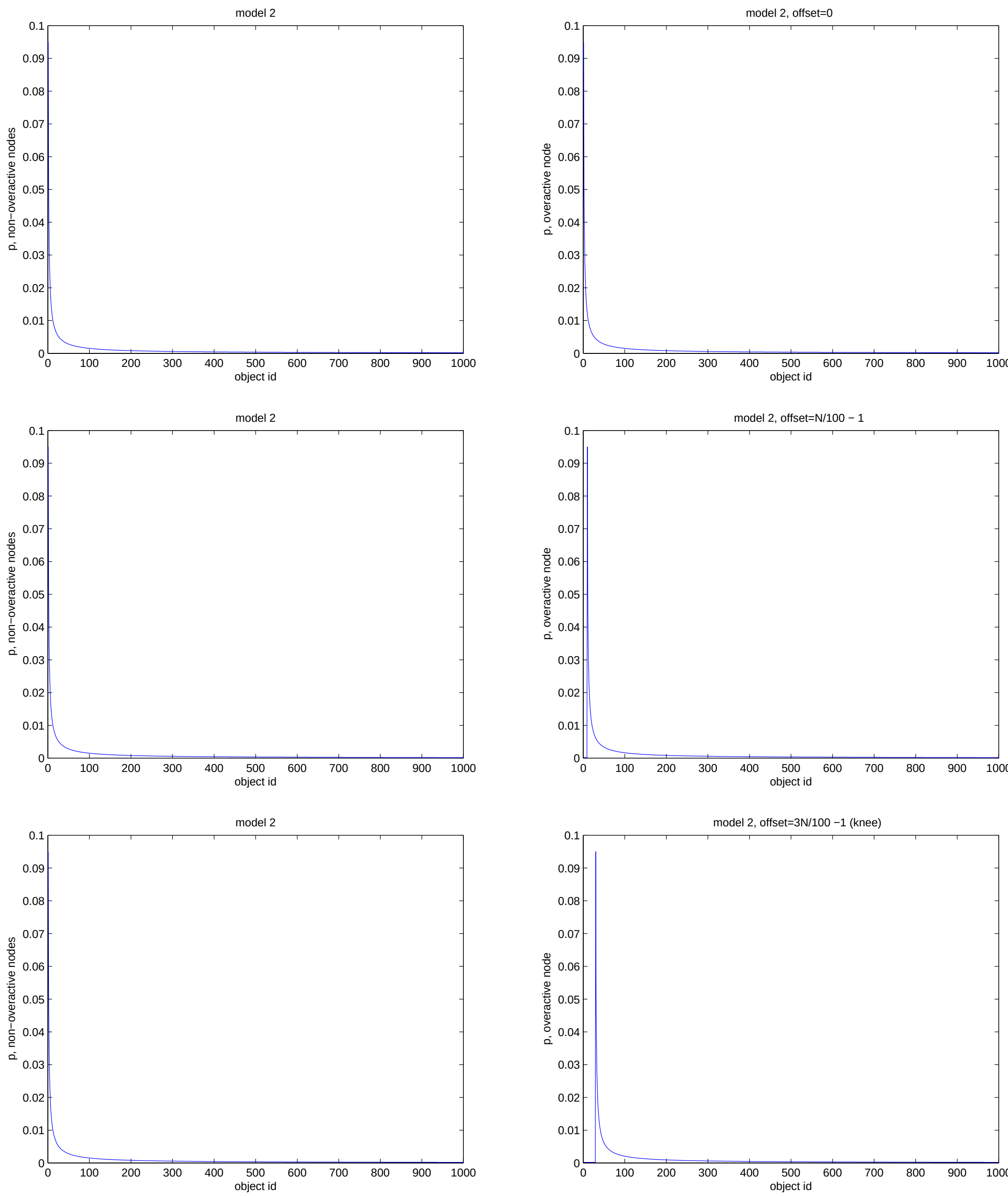


Figure 13: Model 3: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$\{0, N/100 - 1, 3N/100 - 1\}$.

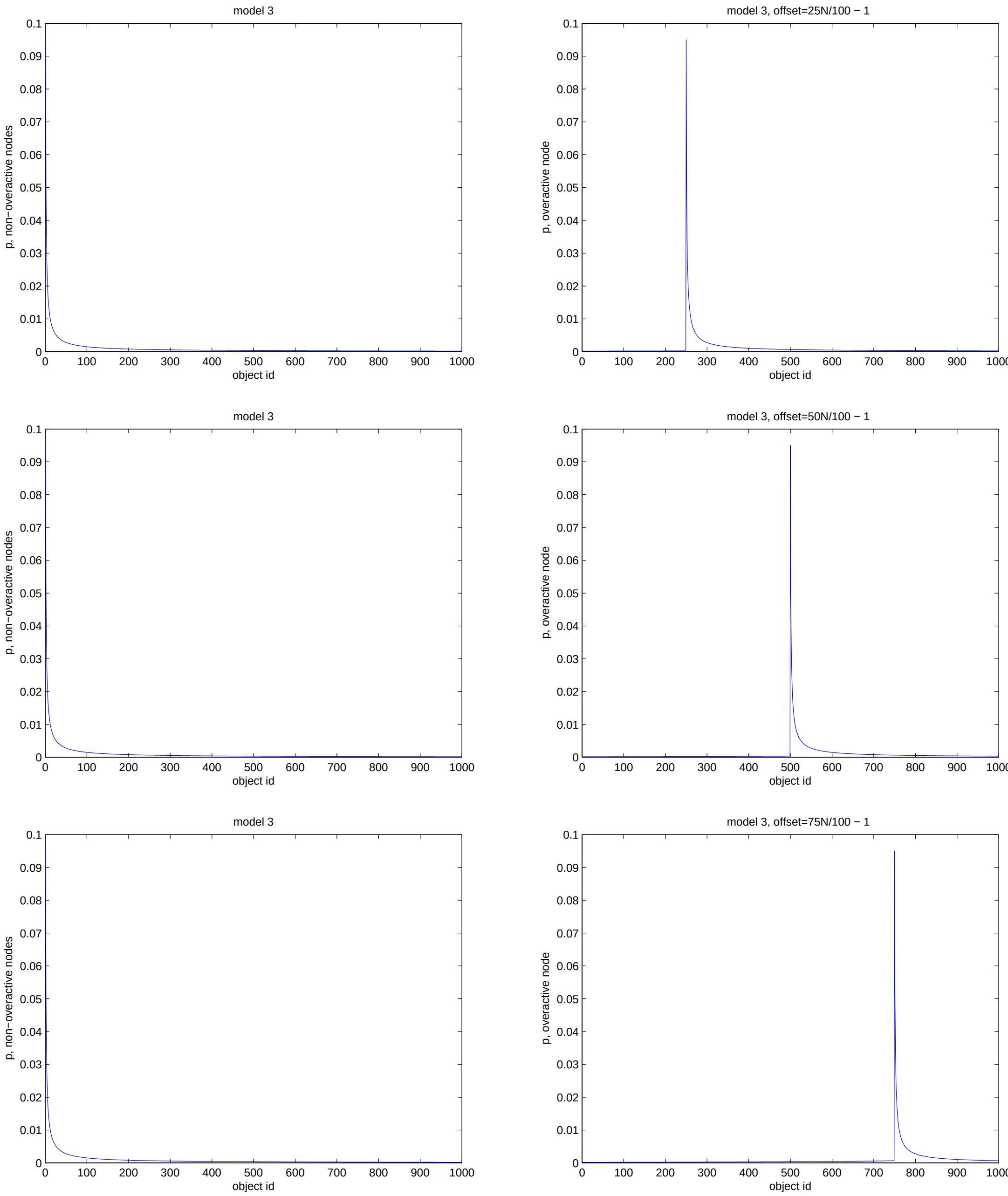


Figure 14: Model 3: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$\{25N/100 - 1, 50N/100 - 1, 75N/100 - 1\}$.

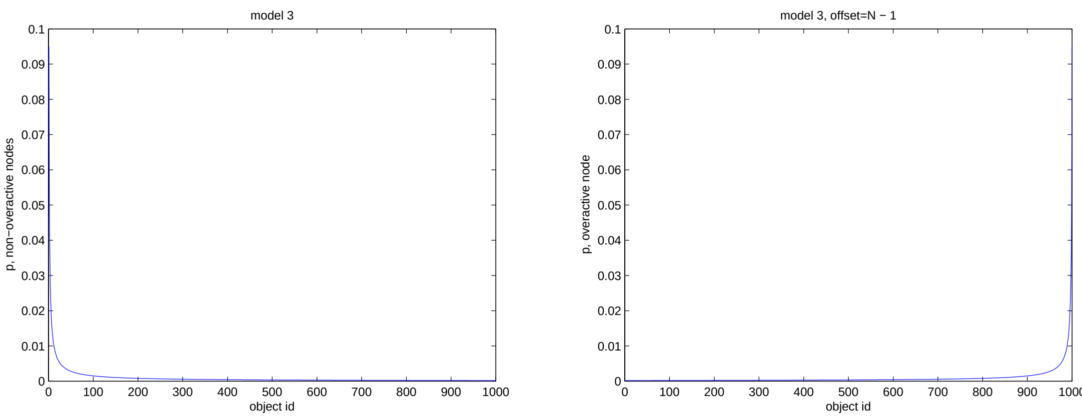
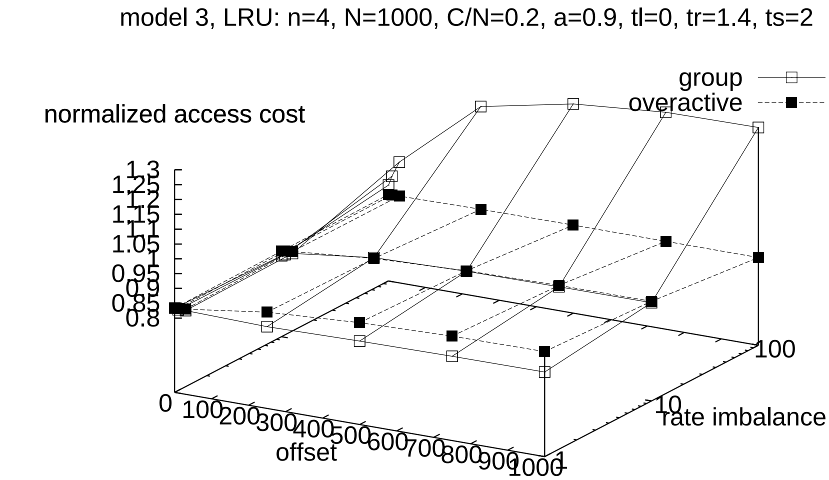


Figure 15: Model 3: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$N - 1$.



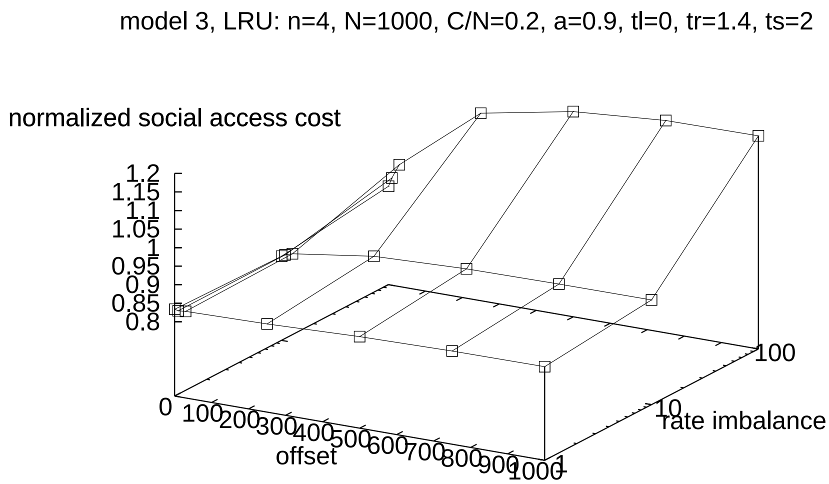
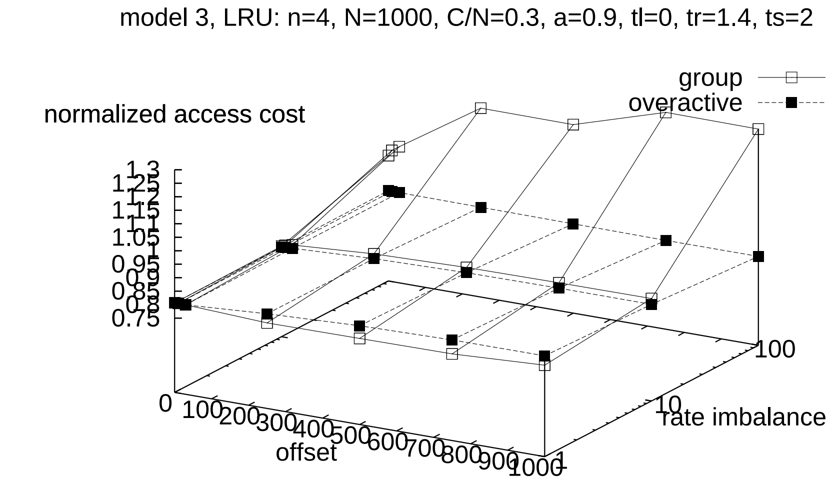


Figure 16: Model 3: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} < N$.



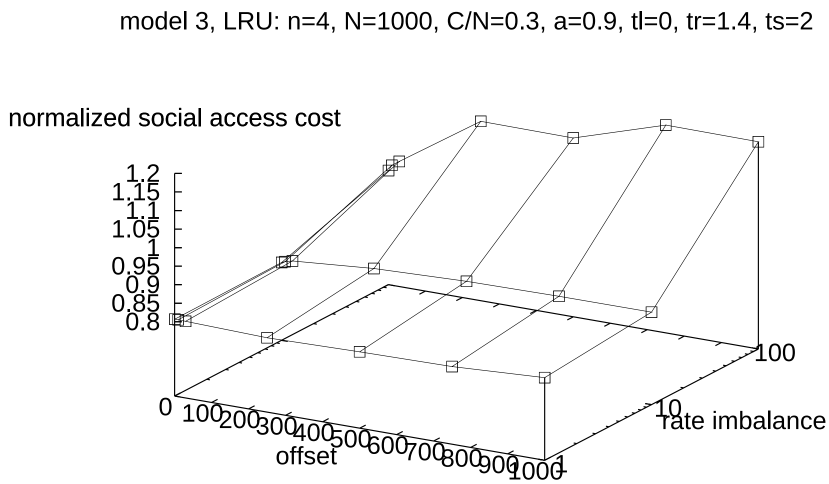
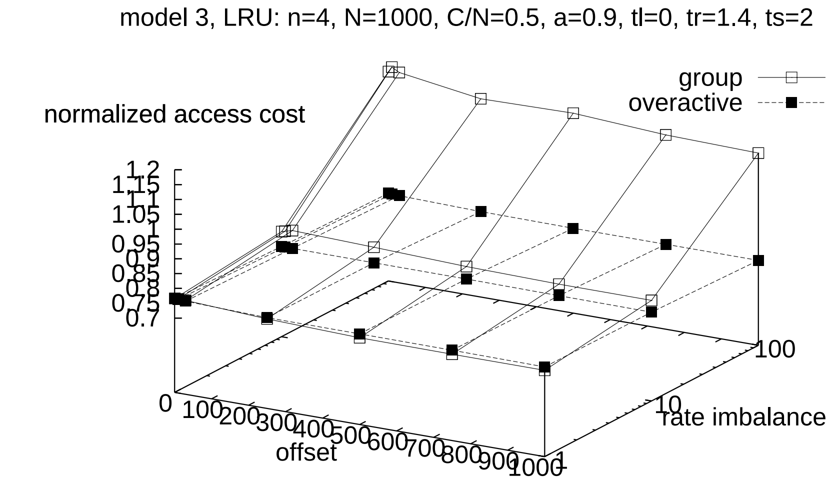


Figure 17: Model 3: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} > N$.



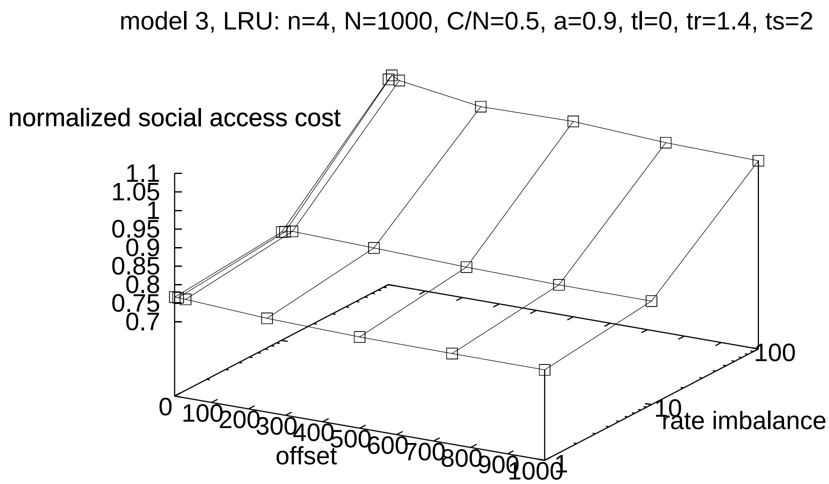


Figure 18: Model 3: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} \gg N$.

4 Model 4

For non-overactive nodes we will maintain the popularity ranking of objects as it was $(o_1, o_2, \dots)$. For the overactive node, however, we will shift the popularity ranking leftwards by an *offset* O, $0 \leq O \leq N$, therefore make object $o_{1+(i-O) \mod N}$ be the ith most popular one. We assign request probabilities taken from the same generalized power-law profile with skewness $a = 0.9$ that is used for the non-overactive nodes.

We plot the demand distributions of the non-overactive and overactive nodes and the individual cost for the overactive and the non-overactive nodes as well as the social cost (the normalization conducted by dividing with the corresponding cost obtained when remote hits are not allowed to affect the local caching state). As it is obvious, mistreatments can occur even under non-homogeneous demand distributions.

Observation 1: When the C_{total} is small (compared to N), the intensity of the mistreatment monotonically decreases with the offset.

Observation 2: When the C_{total} is large (compared to N), the intensity of the mistreatment has a concave profile. The minimum value of the intensity is observed around $N/2$.

Observation 3: For all values of C_{total} we studied the normalized access cost and social access cost is lower than the one achieved when applying Model 1.

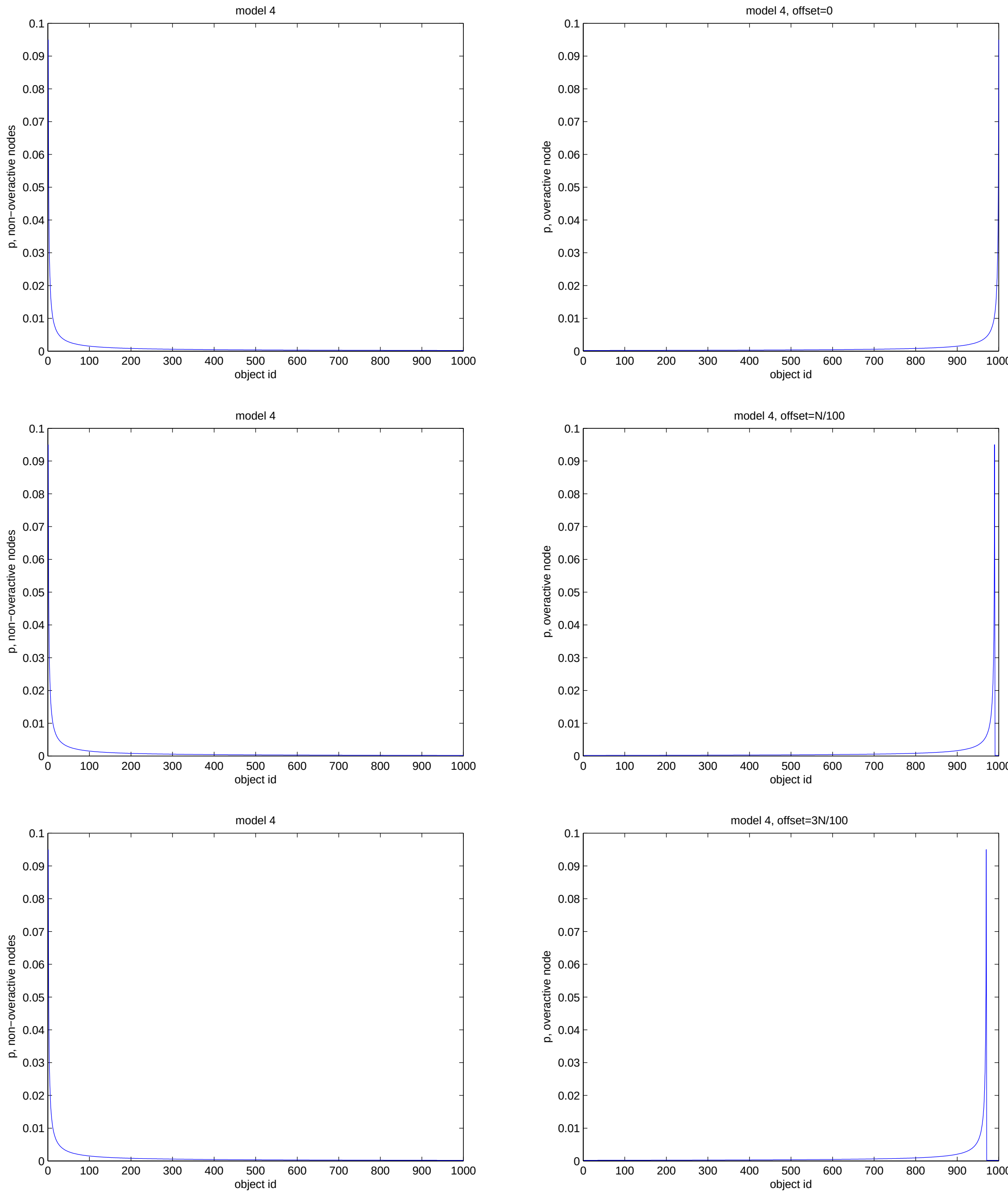


Figure 19: Model 4: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$\{0, N/100, 3N/100\}$.

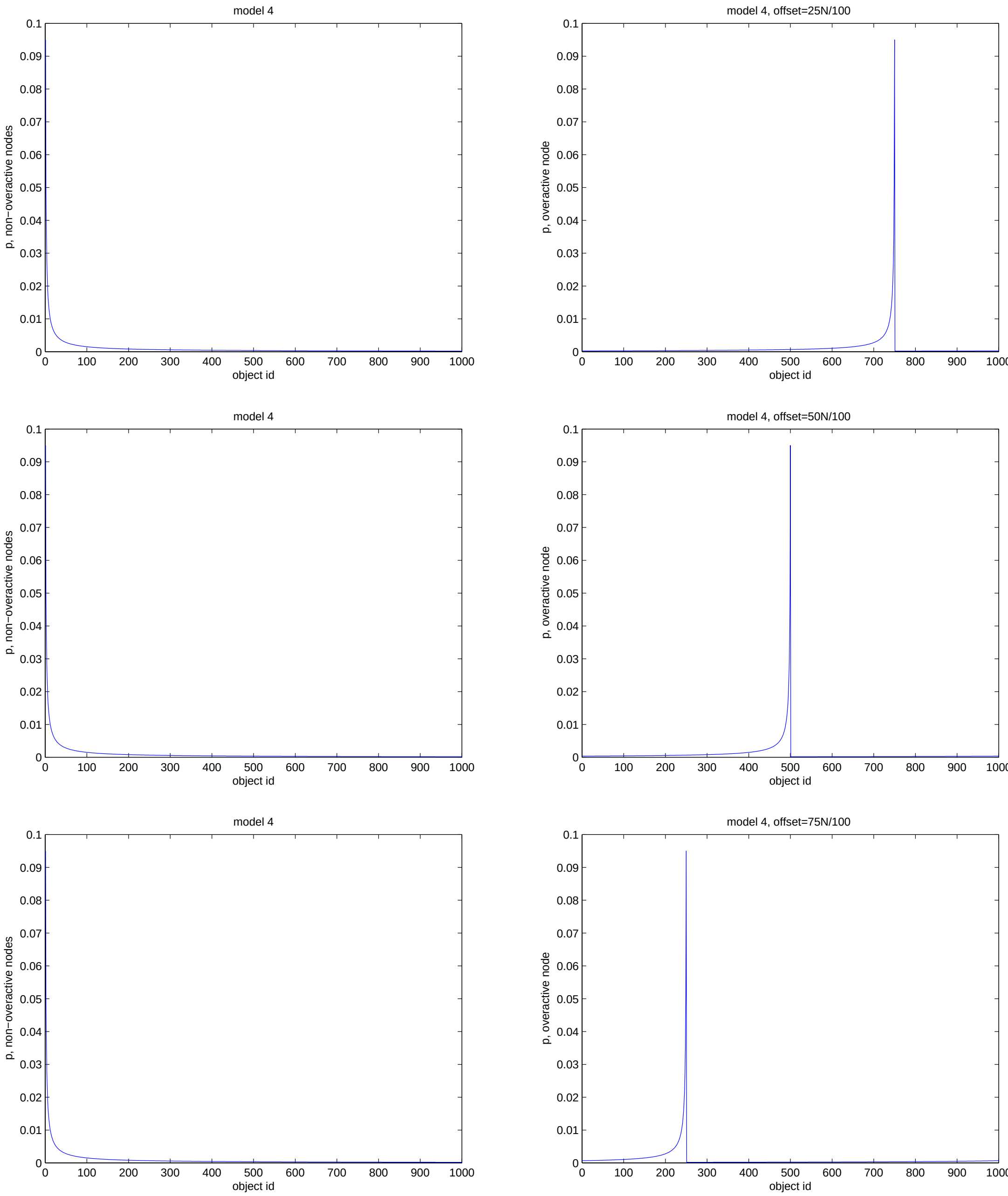


Figure 20: Model 4: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=$\{25N/100, 50N/100, 75N/100\}$.

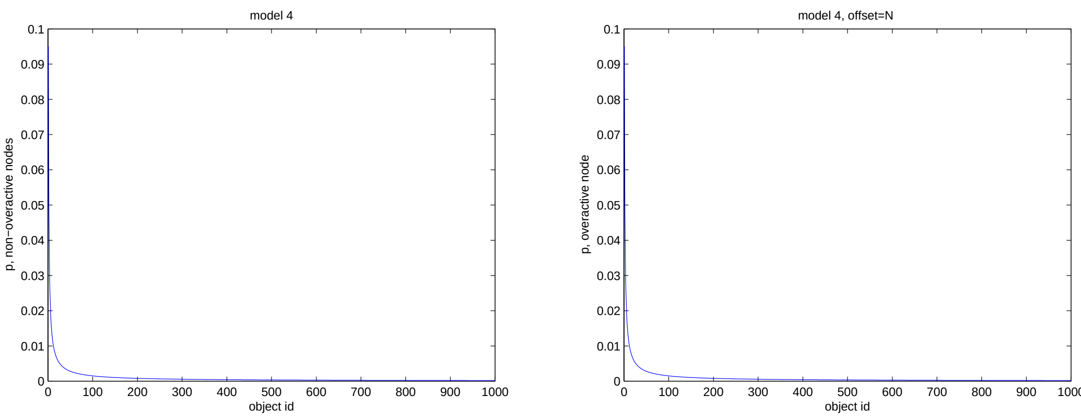
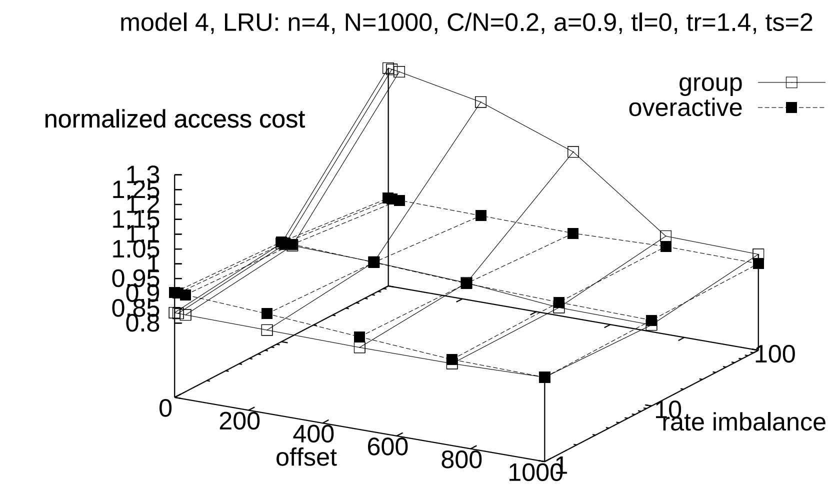


Figure 21: Model 4: Object demand distributions for non-overactive (left) and overactive nodes (right) with offset=N.



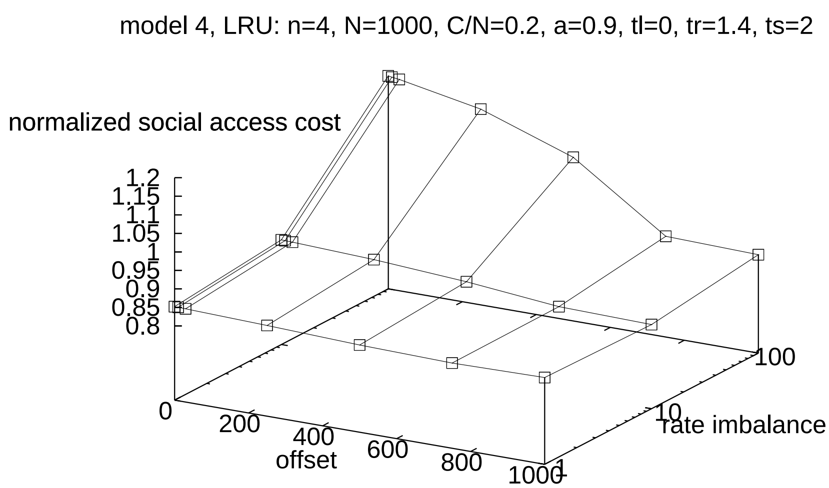
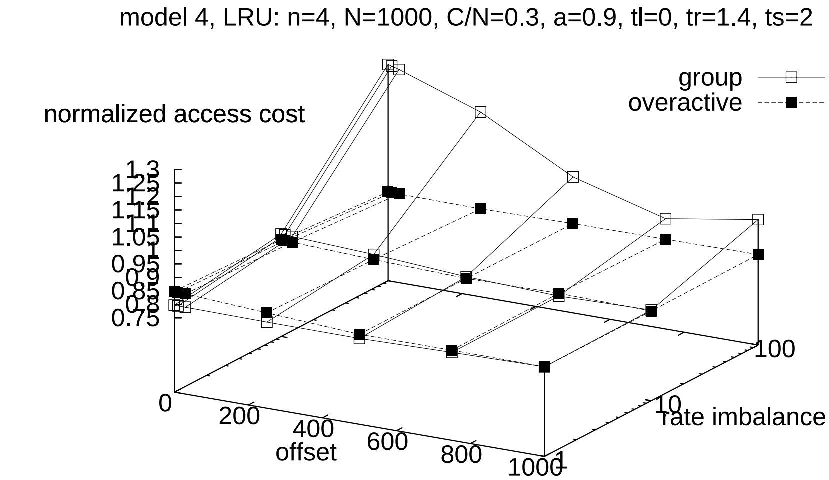


Figure 22: Model 4: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} < N$.



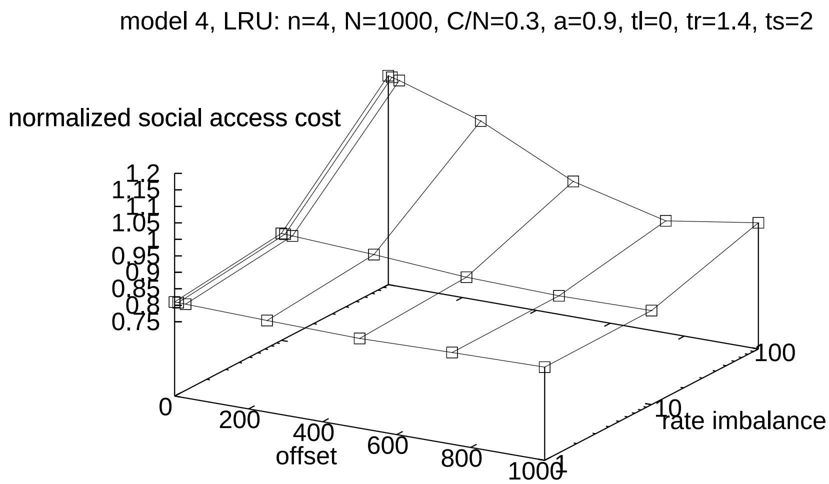
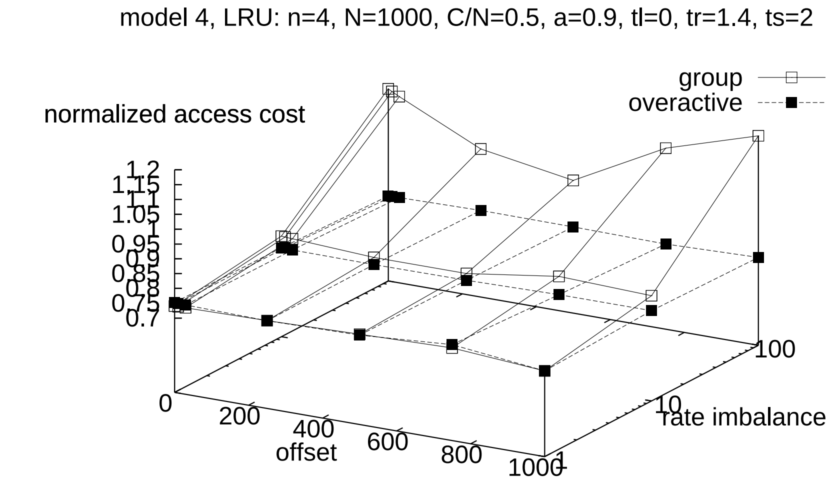


Figure 23: Model 4: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} > N$.



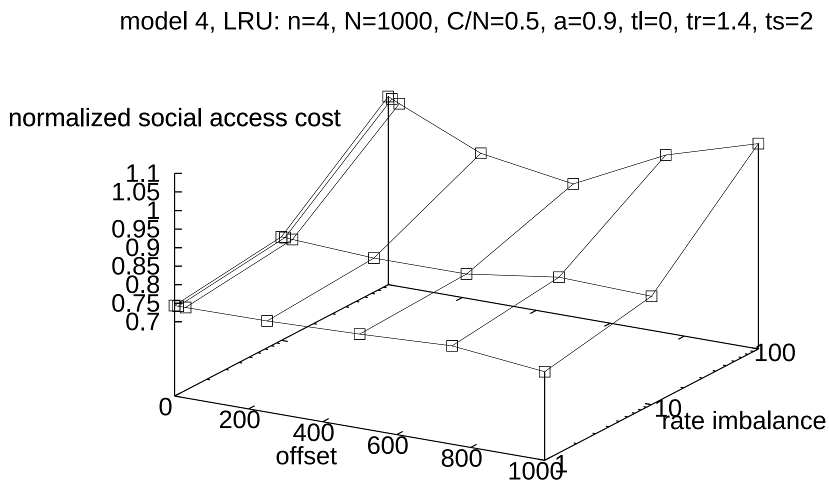


Figure 24: Model 4: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} \gg N$.

5 Model 5

Motivated by applications where ASs contribute caches to a distributed caching group we may assume that although the rank of the objects is similar, the skewness of the demand distribution of non-overactive and overactive nodes may not be identical. We assume that the demand distribution of the group is highly skewed ($a = 0.9$), and the skewness of the demand distribution of the overactive nodes (a_{ov}) differs.

Observation 1: If the C_{total} is small (compared to N), the maximum intensity of the mistreatment is observed when the demand profile of the overactive is uniform and increases with the rate imbalance.

Observation 2: If the C_{total} is small (compared to N), the intensity of the mistreatment has a concave profile (the minimum value is observed around $a_{ov} = 0.5$).

Observation 3: When the C_{total} is large (compared to N), the variation of the intensity of the mistreatment is very small for different values of a_{ov}.

Observation 4: For all values of C_{total} we studied the normalized access cost and social access cost is lowest (compared to models 1-4).

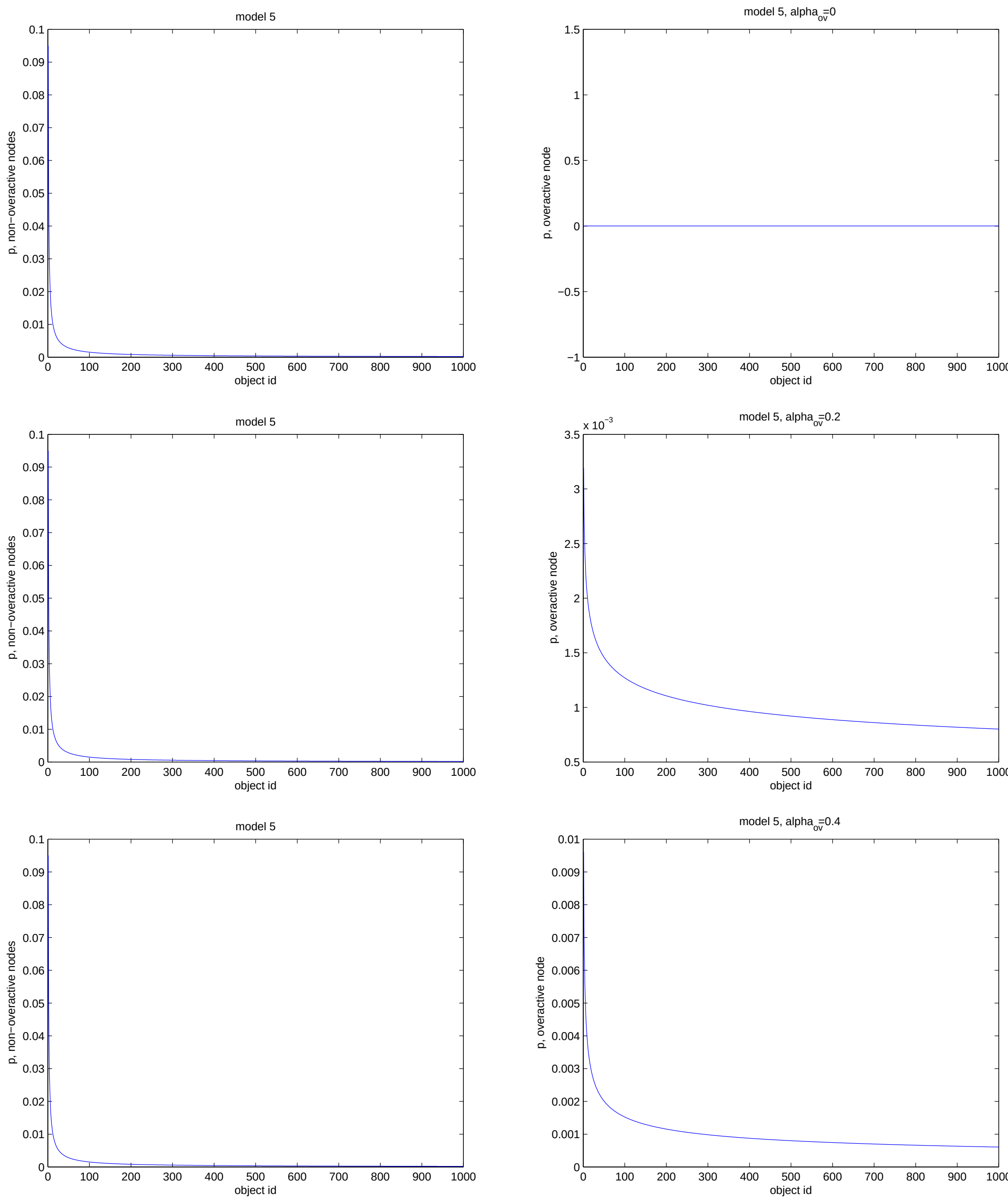


Figure 25: Model 5: Object demand distributions for non-overactive (left) and overactive nodes (right) with $a_{ov} = \{0, 0.2, 0.4\}$.

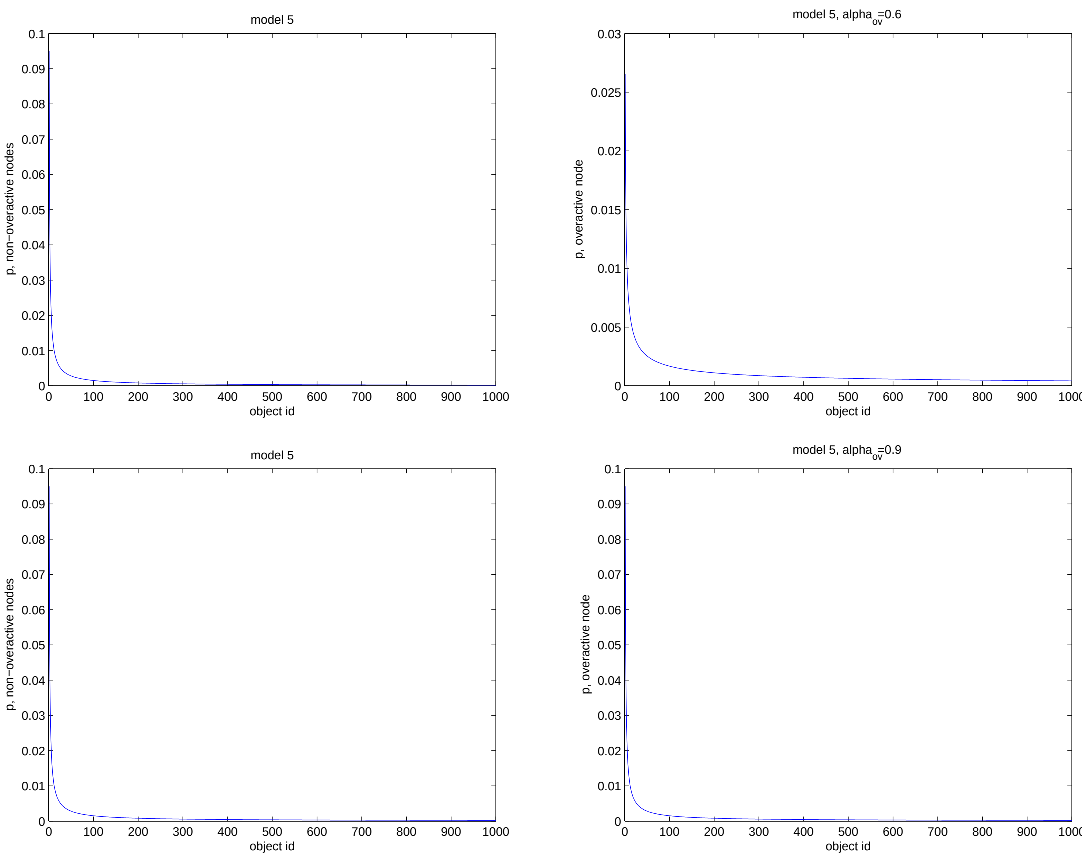
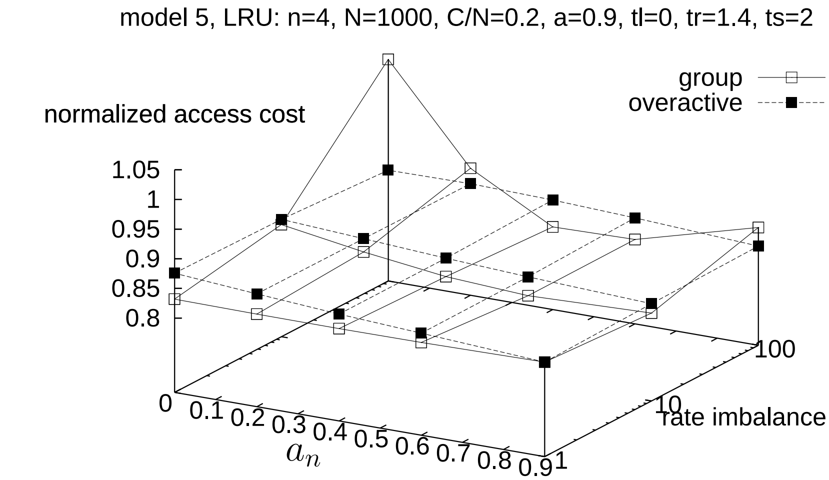


Figure 26: Model 5: Object demand distributions for non-overactive (left) and overactive nodes (right) with $a_{ov} = \{0.6, 0.9\}$.



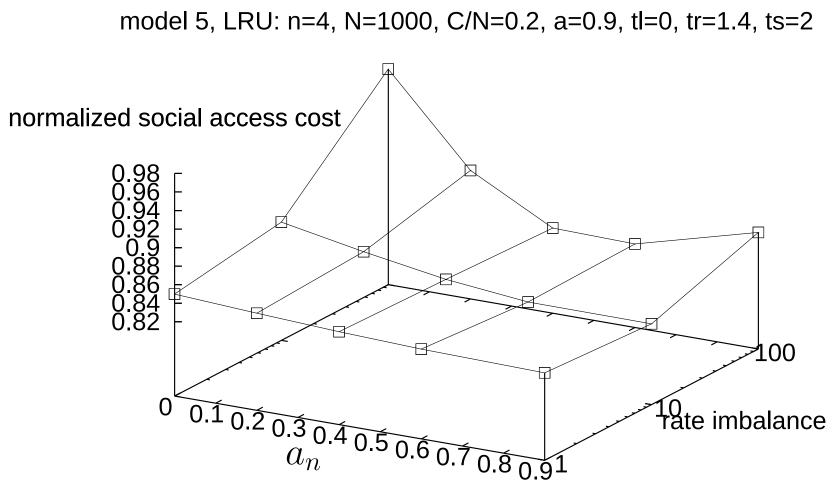
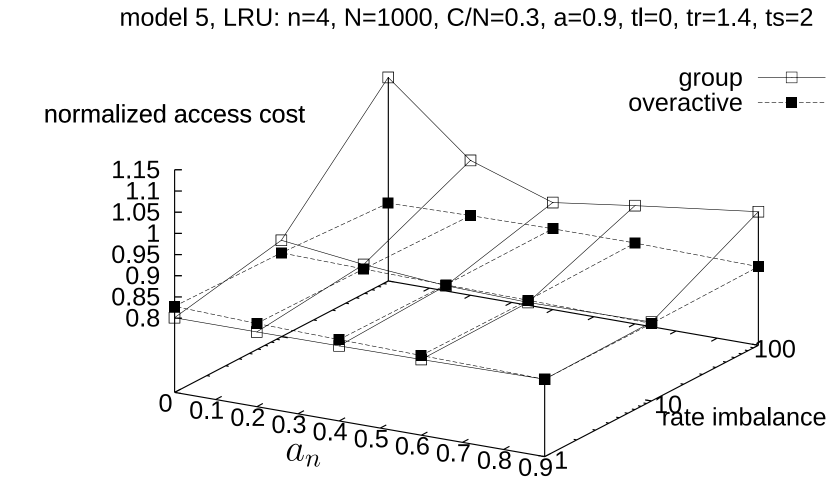


Figure 27: Model 5: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} < N$.



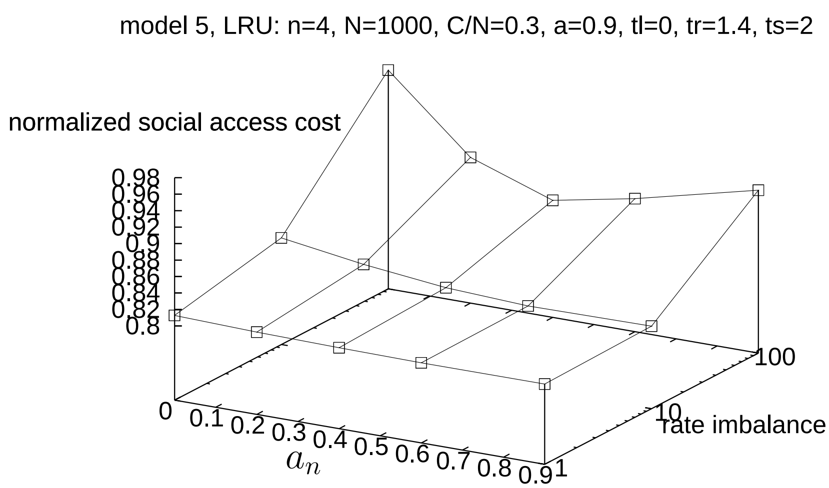
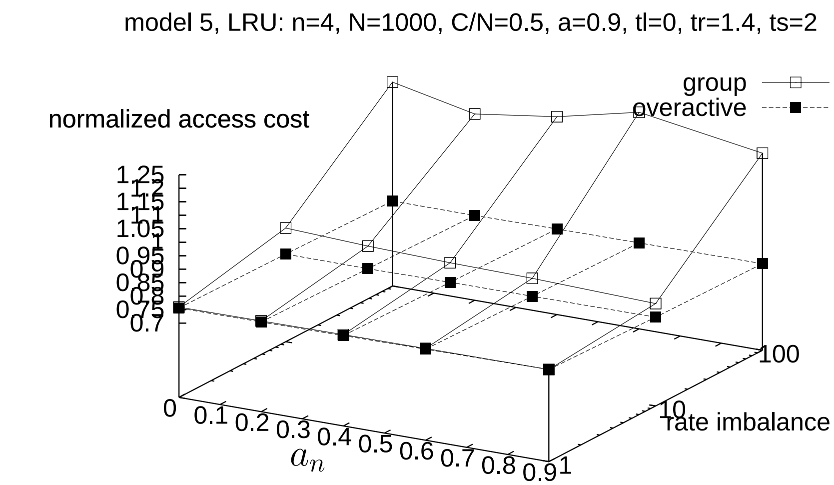


Figure 28: Model 5: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} > N$.



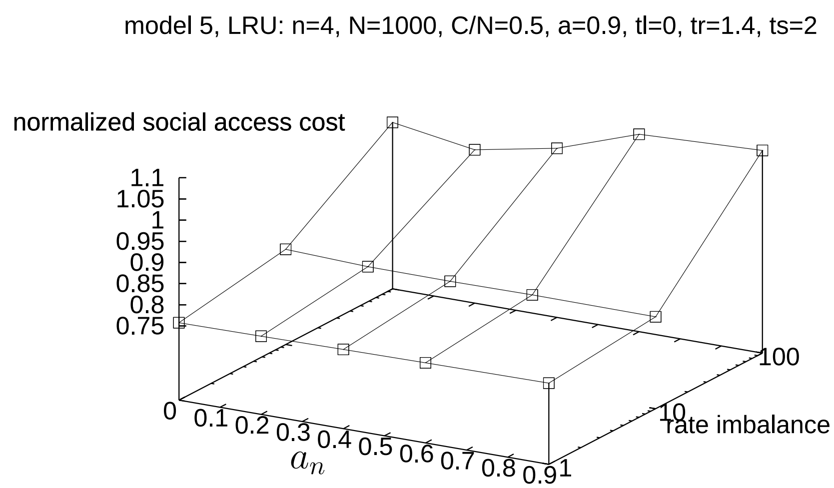


Figure 29: Model 5: Normalized access cost of the overactive and the remaining nodes in the group (top) and normalized social access cost of the nodes in the distributed caching group (bottom) when $C_{total} \gg N$.

References

[1] Nikolaos Laoutaris, Georgios Smaragdakis, Azer Bestavros, Ibrahim Matta, and Ioannis Stavrakakis. Distributed Selfish Caching. Technical Report BUCS-TR-2006-003, CS Department, Boston University, February 7 2006.