

The University of New South Wales
COMP9315 DBMS Implementation
Final Exam

[\[Instructions\]](#) [\[Notes\]](#) [\[PostgreSQL\]](#) [\[C\]](#)
[\[Q1\]](#) [\[Q2\]](#) [\[Q3\]](#) **[\[Q4\]](#)** [\[Q5\]](#) [\[Q6\]](#) [\[Q7\]](#) [\[Q8\]](#)

Question 4 (6 marks)

Consider a table $R(a, b, c, d)$ indexed by tuple-based superimposed codeword signatures. The data and signature files have the following properties:

Page Size	4096 bytes
Number of records (r_R)	10000
Record size (R_R)	32 bytes
Signature size (m)	58 bits
Bits/attribute (k)	10 bits
False match probability (p_F)	1 in 1000

You can make the following (slightly unrealistic) assumptions:

- there are no headers on either data pages or signature pages
each page (except the last) is packed with as many tuples as will fit
- the table has no primary key attribute
- a query like $R(?, b, c, d)$ has exactly 4 genuine matches
- this query also has exactly 10 false matches (consistent with p_F)
- each matching tuple is on a different page
- each false match is on a different page
- no false match occurs on the same page as a genuine match

Based on the above, answer the following:

- How many pages are there in the data file?
- How many pages are there in the signature file?
- How many pages are read in answering the query $R(?, b, c, d)$?
(include both signature pages and data pages, as needed)
- How many pages are read in answering the query $R(?,?,?,?)$?
(include both signature pages and data pages, as needed)
- What feature of the query signature would allow you to optimise the page reads for $R(?, ?, ?, ?)$?
- What would be the values of m and k if we decided to use page-level signatures, rather than tuple-level signatures?

Show all working.

Instructions:

- Type your answer to this question into the file called `q4.txt`
- Submit via: **give cs9315 exam_q4 q4.txt**
or via: Webcms3 > exams > Final Exam > Submit Q4 > Make Submission

End of Question