Feedback Letter

Dear Reviewers and Chairs,

Thank you for allowing a feedback of our research article entitled "Towards Practical and Privacy-Preserving kNN Query over Encrypted Data".

1. Responses to the reviewer #1's comments

O1. The security analysis is not a usual security analysis of encryption schemes. In addition, the problem is a little bit different that standard security problems. For example, the adversary may not need to recover the complete plaintext but a good approximation of it may be enough.

A1 Rev. 1, O1: Thank you for your comments. In our scheme, we prove the security of our scheme in the real/ideal model experiments and intend to show that the view of the adversary in the real and ideal experiments is indistinguishable. That is, the ciphertexts in the real and ideal experiments are indistinguishable. For a clear description, we prove the ciphertexts indistinguishability by showing that both the pattern and the values of the plaintexts are hidden in the ciphertexts. We believe that **ciphertexts are indistinguishable *iff* the pattern and values of the plaintexts are hidden in the ciphertexts.**

The idea of our proof is that

(1) Prove the plaintexts' pattern is hidden in the ciphertexts by reducing the pattern separation problem to the blind signal separation problem. Since the blind signal separation with two Gaussian signals are infeasible, the pattern of the plaintexts with two Gaussian distributions in our scheme is hidden in the ciphertexts.

(2) Prove the plaintexts' values are hidden in the ciphertexts. We show that even if the adversary uses the maximum information it can get, it still cannot recover the plaintexts' values.

For the question that a good approximation can distinguish the real and real experiments, we do not give detailed explanation in the submitted version due to the page limitation. In the following, we show that why a good approximation is not enough to distinguish the real and real experiments. Specifically, Eq. (15) is equivalent to

$$\begin{bmatrix} r_{1,L}\mathbf{x}_1^T & r_{1,L}(-1)_n^T & \mathrm{A}_{1,L}^T \\ r_{2,L}\mathbf{x}_2^T & r_{2,L}(-1)_n^T & \mathrm{A}_{2,L}^T \\ \vdots & \vdots & \vdots \\ r_{p_1',L}\mathbf{x}_{p_1'}^T & r_{p_1',L}(-1)_n^T & \mathrm{A}_{p_1',L}^T \end{bmatrix} \begin{bmatrix} r_j\mathbf{q}_j & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & r_j\mathbf{q}_j & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathrm{A}_j \end{bmatrix} \begin{bmatrix} r_{1,R}(1)_n^T & r_{2,R}(1)_n^T & \cdots & r_{p_1',R}(1)_n^T \\ r_{1,R}\mathbf{x}_1 & r_{2,R}\mathbf{x}_2 & \cdots & r_{p_1',R}\mathbf{x}_{p_1'} \\ \mathrm{A}_{1,R} & \mathrm{A}_{2,R} & \cdots & \mathrm{A}_{p_1',R} \end{bmatrix} = \begin{bmatrix} C_{1,j,1} & C_{1,j,2} & \cdots & C_{1,j,p_1'} \\ C_{2,j,1} & C_{2,j,2} & \cdots & C_{2,j,p_1'} \\ \vdots & \vdots & \ddots & \vdots \\ C_{p_1',j,1} & C_{p_2',j,2} & \cdots & C_{p_1',j,p_1'} \end{bmatrix}$$

We denote the above equality as LQR = C. First, we can choose two random matrices D_1 and D_2 satisfying the first n diagonal elements are same as the second n diagonal elements. Then, we can easily find two matrices X_1 and X_2 such that $\mathrm{X}_1\mathrm{L} = \mathrm{D}_1$ and $\mathrm{R}\mathrm{X}_2 = \mathrm{D}_2$. In this case, we have $Q = (\mathrm{X}_1\mathrm{L})^{-1}\mathrm{X}_1\mathrm{C}\mathrm{X}_2(\mathrm{R}\mathrm{X}_2)^{-1}$. For any random

D_1 and D_2, we can obtain a Q. That is, Eq. (15) has an infinite number of solutions. Even if the adversary obtains some approximate solutions, the probability that the approximate solution is $\mathbf{q}_j$ is negligible.

O2. The paper is incremental and improves an existing idea although the approach seems valid and correct.

A2 Rev. 1, O2: Our work follows Wong et al. work [19] (published in Sigmod 2009). After Wong et al. proposed the first matrix-based encryption kNN query scheme, the idea of matrix encryption is widely used to preserve the privacy of queries. However, the security of Wong et al.'s scheme is not KPA secure. In this work, we present KPA-secure matrix encryption, which is the first matrix encryption with high efficiency and KPA security.

If a scheme is insecure, it definitely can not be deployed in the real scenarios. Therefore, it is not a trivial work to make an insecure scheme into a secure one, especially for a scheme was found insecure, but no improvement in the past then years.

O3. The authors do not discuss well how the proposed approach will address recent attacks on database queries that use the answers over different queries over time to infer the real or approximate values of plaintext. See for example the papers on access pattern attacks and volume based attacks.

A3 Rev. 1, O3: Thank you for your valuable comments. We have realized the access pattern privacy is a critical issue in database queries. However, due to the page limitation, it is not our research focus in this paper, we We do not will consider the access pattern privacy in our future work. In this scheme, we only focus on achieving the KPA-secure data privacy.

O4. The proposed approach for k-NN is based on linear scan. If the database size is large and the system has to read the database from the disc, it may be better (and faster) to use an interactive k-NN scheme with small number of interactions.

A4 Rev. 1, O4: In this paper, we only give a simple scenario to show the efficiency of our scheme. In fact, our AME scheme is a basic scheme, which can be integrated with lots of mature data structure techniques to improve the kNN query efficiency, e.g., kd-tree, M-tree etc.

2. Responses to the reviewer #2's comments

O1. The core technical contribution here is a novel property-preserving encryption scheme, which is a cryptographic contribution rather than a data processing contribution. Accordingly, I would suggest that the paper will receive better feedback from the cryptographic community.

A1 Rev. 2, O1: Our work follows Wong et al. work [19] (published in Sigmod 2009) and focuses on database query privacy. Thus, we believe that our work falls within the scope of the Sigmod conference.

Q2. ***I was not convinced by your security claim. For instance, you have the following passage in your proof: "Since each of $R_{i,L}$, R_j and $R_{i,R}$ has at least two dimensions following the Gaussian distribution, it is impossible for $\mathcal{A}$ to separate the patterns of $\{\{\mathbf{x}_i\}_{i=1}^{p_1}, \{\mathbf{q}_j\}_{j=1}^{p_2}\}$ from their corresponding ciphertexts. This is because when there are two Gaussian signals, there will be an infinite number of patterns of $\{\mathbf{x}_i\}_{i=1}^{p_1}$ and satisfying the ciphertexts. Thus, it is hard for $\mathcal{A}$ to distinguish them. In this case, $\mathcal{A}$ cannot find the patterns of the plaintext database records and query records from their corresponding ciphertexts. Thus, $\mathcal{A}$ cannot distinguish the $\mathrm{View}_{\mathcal{A},\mathrm{Real}}$ from random ciphertexts based on the patterns of the database records and query records." The language above is rather informal and unconvincing. Just because there are infinite number of patterns satisfying the ciphertexts does not mean that A cannot distinguish the real view from random ciphertexts. Even if A can learn one bit of information in addition to the permitted leakage, then your security claims fall apart.***

A2 Rev. 2, O2: In this passage, we aim to explain that the pattern of the plaintexts can be hidden in the ciphertexts.

3. Responses to the reviewer #3's comments

O1. ***I do not find the security analysis of the proposed AME scheme sufficiently convincing. The security definition based on the Real/Ideal paradigm is appropriate, and the construction of the simulator for the Ideal experiment seems to be sound. Nonetheless, I believe that this paradigm is used improperly in the paper, as the proof which shows that the scheme meets this security definition seems to be focused only on the security vulnerabilities affecting the existing ASPE schemes described in the preliminaries, i.e., the known plaintext attack strategies and the ciphertext-only techniques that permits to discover the distribution of the plaintext values. Unfortunately, proving that these attacks are not possible is not sufficient to meet the security definition based on the Real/Ideal paradigm, as this definition mandates to prove that the views of the two experiments are computationally indistinguishable for any polynomial-time adversary. The security analysis should thus focus on arguing why the ciphertexts employed in the real protocol are indistinguishable from the random ones constructed by the simulator in the Ideal experiment, rather than discussing which information about the plaintext values can be learned from the ciphertexts observed in the Real experiment. Indeed, employing a security definition based on the Real/Ideal paradigm has the advantage that the proof can simply focus on proving the indistinguishability of the two experiments, as this imply that the real protocol allows the adversary to infer no more information than the leakage employed in the proof, in turn removing the need for the prover to consider all the possible attack techniques available. In conclusion, the authors should complete the security proof by proving such indistinguishability to convince the reader that the proposed AME scheme is secure. For instance, it is not clear whether the ciphertexts employed in the two experiments remain indistinguishable in case the adversary chooses plaintext records with all zero entries;***

the authors should provide a general argument that shows the indistinguishability of the ciphertexts in the two experiments for any choice of plaintext values.

A1 Rev. 3, O1: In our scheme, we prove the security of our scheme in the real/ideal model experiments and intend to show that the view of the adversary in the real and ideal experiments is indistinguishable. That is, the ciphertexts in the real and ideal experiments are indistinguishable. For a clear description, we prove the ciphertexts indistinguishability by showing that both the pattern and the values of the plaintexts are hidden in the ciphertexts. We believe that ciphertexts are indistinguishable *iff* the pattern and the values of the plaintexts are hidden in the ciphertexts.

The idea of our proof is that

(1) Prove the plaintexts' pattern is hidden in the ciphertexts by reducing the pattern separation problem to the problem of blind signal problems. Since the blind signal separation with two Gaussian signals are infeasible, the plaintexts' pattern is hidden in the ciphertexts.

(2) Prove the plaintexts' values are hidden in the ciphertexts. We show that even if the adversary uses the maximum information it can get, it still cannot recover the plaintexts' values.

As our proof, our scheme can hide the pattern and plaintexts of the ciphertexts. Thus, the adversary still cannot distinguish the real and ideal experiments even all plaintexts are zero entries.

O2. The security analysis of the k-NN query scheme, reported in Section 5.2, should explicitly state that the proposed solution does not guarantee the access pattern privacy of the queries, as the adversary trivially observes which records are the k nearest to the queried one, although their content is properly concealed thanks to the privacy guarantees on query results reported in the security analysis.

A2 Rev. 3, O2: Our scheme indeed cannot preserve the privacy of access patterns.

O3. Although the experimental evaluation thoroughly validates the asymptotic trend in the performance of the proposed solution and the existing ones, it lacks two important additional evaluations. First, instead of employing an artificial dataset, the practicality of the approach should be assessed by evaluating the performance, especially the response time of the queries, of the proposed solution on a real-world use case. In particular, the authors should consider real-world datasets with millions of records, which seem to be in the realm of practicality for the proposed solution given the execution times reported in the experimental evaluation. In addition, it would be interesting to observe the performance overhead of the proposed k-NN query scheme w.r.t. a baseline solution with no security guarantees, which performs k-NN queries over plaintext data. This evaluation may provide to application designers an estimation of the security cost of the proposed solution.

A3 Rev. 3, O3: Since our scheme performs kNN query by linear scan, so its efficiency is not affected by the distribution or skewness of the dataset. Thus, the performance of our scheme on an artificial dataset is the same as that on a real dataset.

In the performance evaluation, we compare our scheme with the ASPE scheme. In fact, the performance of the ASPE scheme is the same as the baseline solution with no security guarantee. This is because the ASPE scheme encrypts a vector into a vector and the security does not incur any additional computational cost. In this case, comparing our scheme with the ASPE scheme is equivalent to compare it with the baseline solution.

O4. The authors should clarify why they deem as costly the setup of a secure channel between users and the cloud provider: indeed, given that a cloud provider is likely to already own a digital certificate for its servers, establishing a TLS based secure channel between each user issuing a query and the cloud server should be straightforward. Hence, it is unclear why the authors decide to employ different encryption matrices for each user instead of relying on a secure channel, also considering that the solution adopted by the authors has even weaken privacy guarantees than a secure channel. Indeed, since the results of the queries are deterministically encrypted, it is possible for any eavesdropper to determine if the set of k nearest records retrieved by distinct queries of the same user are disjoint or not, an information which would be hidden in case these results are sent through a TLS based secure channel. The authors should motivate the advantages given by their solutions w.r.t. a standard secure communication channel.

A4 Rev. 3, O4: We agree with you and we will remove the part that distributing different users with different matrices in the final version.

O5. The manuscript is in general clear and well-structured. However, there are some improvements that may ease the comprehension of the paper. First of all, the concept of matrix encryption scheme is never explained, not even informally. It is mentioned in the introduction as a building block of the privacy-preserving k-NN query solution, and thus this concept should be properly framed already in the introduction and possibly also formally defined in the preliminaries. In addition, the description of the COA technique in Section 3.1 employs notions which are probably clear in the signal processing domain but that need to be clarified for a general audience. For instance, what is the Independent Component Analysis? What does it mean to separate source signals? The authors should briefly discuss these notions or they should at least provide proper references that allow the interested reader to understand these concepts.

A5 We will discuss these concept clearly in the final version.

O6. The results of the experiments reported in the experimental evaluation are not sufficiently commented. Indeed, Section 6.2 mostly describes the trends and the performance gap between the different solutions that can be observed in the figures reporting the results of the experiment, but with no further discussion on the described

results. For instance, in some experiments, the trend observed in the experimental results differs from the one expected according to Table 1 (e.g., the query processing time depends linearly on the dimension d for the proposed scheme instead of quadratically as in Table 1); the authors do not try to motivate this discrepancy. I suggest to enrich the discussion of experiments with some comments on the results and on the sources of the performance gap observed between their solution and the existing ones.

A6 In Fig. 2(e), the computational cost of query processing linearly increases with d while the computational cost is $(2d + 9)(N + \alpha * (2d + 9) + \alpha * log_2 k)$. The reason for this case is that α (i.e., the number of data records need to be inserted into the heap) is small. Thus, the main computational cost of query processing is $(2d + 9) * N$, so it linearly increases with d. We will explain these figures clearly in the final version.

In my opinion, authors should briefly explain why matrices over real numbers are employed in their AME scheme instead of relying on integer arithmetics, which is more efficient. I suspect that the adoption of real numbers is required for the security of the scheme, but this should be explicitly motivated.

Rev. 3, Additional Remark: The reason that we employ the matrix encryption over real number is that the real numbers is more secure that integers. For example, real numbers do not have the integer factorization problems. We will explicitly stress this motivation in the final version.

4. Responses to Editor

Reviewer 1 has 4 remarks and most of them are constructive. But one is not inline with reviewer instructions as follows.

(1) Review 1 Remark O2 is not inline with reviewer instructions as it refers to our paper is incremental. Although the problem of our scheme is an existing problem but the solution is a new one.

Reviewer 2: Reviewer 2 has two remarks, but both of them are not inline with reviewer instructions as follows.

(1) Review 2 Remark O1 is not inline with reviewer instructions as it refers to our scheme is a property-preserving encryption scheme and should be submitted to the cryptographic community. Our scheme focuses on the kNN query privacy over encrypted database, and the previous work, i.e., Wong et al. work [19], also published in Sigmod 2009. Thus, we think that our scheme falls in the scope of Sigmod.

(2) Review 2 Remark O2 is not inline with reviewer instructions as it refers to our security analysis is not convincing by showing that some description is not informal.

Reviewer 3: The comment from Review 3 is really constructive.