Secure Multi-Party Computation for Sorting Alphabetic Words

Arpesh Singh, Vishal Passricha
Computer Engineering Department, National Institute of Technology
Kurukshetra, Haryana, India

ABSTRACT-Internet is growing very fast and computation over the internet is very much common. Some tasks are computed jointly between multiple parties, where people supply their corresponding inputs to be computed mutually with the inputs of other parties to obtain the output. Sometimes these computations are occurred between mutually untrusted parties or say, competitors, hence the parties don't want toshare their data to other parties. To perform such kind of computation, Secure Multiparty Computation (SMC) comes into view. SMC ease the problem of data sharing by parties as computation is securely done over secure inputs (encrypted inputs) provided by different parties and generate the corresponding output in a secure form which is correspondingly delivered to each party. However, there is no proper solution for sorting such words. To solve this problem, we proposed an algorithm that generate a cipher value of a word by using polynomial function and thesecipher values are used in sorting problem without decrypting the word. The proposed algorithm is less complex and offers high data security.

KEYWORDS-*Alphabetic Sorting, Encryption Technique, Polynomial Function, Privacy, Secure Multiparty Computation.*

INTRODUCTION

The increasing use of internet makes the people dependent upon the cyber world and tends all the operations over the internet. The huge amount of dependency over the internet has also generated option of co-operative computation, where people jointly conduct their tasks of computation by providing their respective inputs. In co-operative computation, two or more than two parties provide their respective inputs to obtain a result which is computed using all the data provided.However, in the present era, misuse of data is a big issue; whichpreventscomputing parties to share data. To overcome this problem, secure multiparty computation (SMC) comes in view where data is computed in the encrypted form itself without knowing the original data provided by the computing parties.

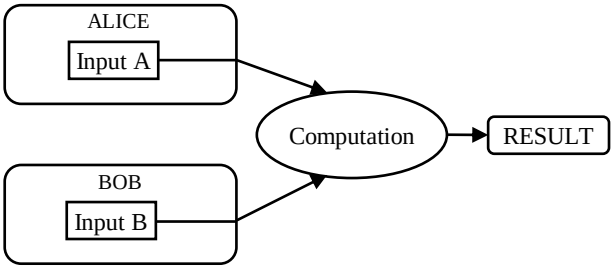

Fig 1: Multiparty Computation Model

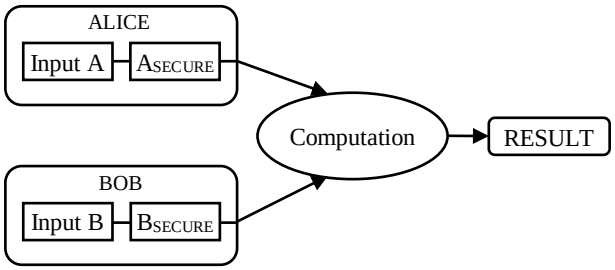

Fig 2: Secure Multiparty Computation Model

In SMC, the inputs are provided in encrypted form by the computing parties.After collecting the inputs,the server applies function directly to the encrypted values and generates encrypted output which is sent back to users. In SMC, the input is encrypted in such a way that it remains secure and it is also computed without decryptionby maintaining the correctness of output[1].

There are several problems of secure multiparty computation like database query[2], geometric computation[3], scientific computation[4], data sorting and selection[5].

A) Database Query

SMC, the major problem is the matching of query in database, where one party needs to check his input in other's database, maintaining the privacy of both the parties. In this problem, matching is either given as perfect 'yes' or 'no' or the output may be produced in the form of correlation score which denotes the percentage of matching with the like items in the database.

Example: Let's say Alice wants to search some entry to be present in Bob's database, but she doesn't want to share details with Bob because of personal information in the input. Bob also does not want to give his whole database to Alice to search her query.

B) Geometric Computation

In geometric computation, points are provided as inputthen different geometric functions are applied on them to obtain the corresponding result. The problem in co-operative geometric computation is the privacy of the coordinates provided as input.

Example: Two persons standing at point P_1 and P_2 want to know their distance from each other but don't want to reveal their original location for a security purpose.

C) Scientific Computation

Linear equation problems are solved mutually by using a number of linear equations provided by different parties without revealing the actual equation to other party. Such problems occur between two competitors working on same project proving individual linear equations.

Example: Alice has some private linear equations say $M_1x = a$ and Bob also has private equations say $M_2x = b$. They mutually want to compute the value of x that satisfy both.

D) Data Sorting and Selection

Data sorting of private dataset provided by two or more parties to sort the data and server returns the indices of their respective inputs from the combined sorted dataset. Further, this will help in finding k^{th} element or median of the dataset obtained. Comparison of the dataset is needed using SMC in this problem.

Example: Let Alice has some private dataset D_1 and Bob has private dataset D_2. They mutually want to sort the dataset and need to know the k^{th} element belongs to which party.

In this paper, we have considered the problem of sorting the words provided by the different user in alphabetic order. A polynomial algorithm is used to convert the word into correspondingcipher value in such a way that the server can apply the sorting function without decrypting the value into word and provides the sorting order to the parties as output.

RELATED WORK

Secure Multiparty Computation was introduced in 1982 by A. Yao [6]and was further extended by O. Goldreich et al.[7]. The computationproblem is first represented as a combinatorial circuit, then the parties run a short protocol for every gate in the circuit. This size depends on the size of the input domain, and on the complexity of expressing such a computation.

According to literature, mostly SMC research are theoretical studies and few implementing problems have been studied. Several examples that are using SMC for secure input computationareinformation retrieval problem, privacy-preserving statistical database, and privacy-preserving data mining[8][9][10][11].

A. Shamir [12]proposed a technique of sharing the data in the form of polynomials.In his work, the data is distributed into '*n*' number of the polynomial equation having similar data but cannot be decrypted without knowing more than (*n*/2) equations. Polynomials can be used as encrypted value operations because the value of data is changed according to the polynomial equation which is easy to use in comparison problems.

There are some secure multiparty computations for comparison of number using bit decomposition protocol [13].Further some improvement was done in sorting numbers without using bit decomposition protocol[14].

SMC problems exist in some more domains other than the problem described above, most of them are not yet solved. Recent problemsemerge if we combine privacy with co-operative computation.

PROPOSED SCHEME

There is no proper algorithm for word sorting in secure multiparty computation. Earlier, character by character each word is changed into ASCII value and then computed using bit decomposing comparison algorithm which makes the solution very much complex, large and even time-consuming. In this paper, authors proposed an algorithm which works on the whole word instead of character by character approach.

The newly proposed algorithmencrypts the English words in such a way that it generates a unique cipher value which is used as input data for sorting the provided words in dictionary order (alphabetic order). Weight is assigned to each alphabet (i.e. A = 1, B= 2, …., Z=26) to compute polynomial function. In short, the proposed algorithm usesthe wholeword instead of characters which makes it less complex and time efficient.

A) Setup

All the user shares a number 'n' and an 'α' set (set of distinct variables which are used as multipliers) of 'n' number of elements, say $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n\}$secretly.

To achieve better results and maintaining security, we applied a constraint on 'n', that it should be greater than the maximum length of the word to be computed and the value of all elements in α set should be greater than 26 and must follow decreasing order i.e. $\alpha_i > \alpha_{i+1}$ for all i in α set.

B) Encryption

Each party encrypts their word by using polynomial function providing 'n' and α set to generate their respective cipher value which is shared to the server for further

computation. Sharing to the server is done using public key encryption techniques like El-Gamal encryption technique[15].

The proposed function to encrypt the word into cipher value is asfollows:

$$C = \sum_{i=1}^{n} \alpha_i^{n-i} w_i$$

where C is the calculated cipher value and w_i denotes the weight of i^{th} character of the word.

C) Sorting

Third party server collects the values from all the users and compares the cipher value. The above-mentionedpolynomial function generates the cipher value using the weight of alphabets in such a waythat the cipher value of word occurring first in dictionary order is always less than the cipher value of word occurring later. Therefore,the server just sorts the cipher value in ascending order and assign a rank(index) to each cipher value. Each user receives back index value of their corresponding cipher value in return from server.

After sorting the cipher value server returns the corresponding index value to eachuser describing the position of their respective word in the collective dataset provided by the users.The outputis also sent in the encrypted form hiding the position ofthe word of one user from the other.

RESULT&CONCLUSION

The proposed scheme generates the result for all the words in the dictionary and has given correct output for all different pairs of words taken in test cases.

The proposed scheme is secure for every communication that is taking place in multiparty computation and the privacy of data remains secure from all the other users involved in the computation.

A) Security Analysis

The proposed algorithm computes the cipher value of the given word using a polynomial function with the exponential power of 'n' that makes the decryption harder for the attacker to decrypt the cipher value into the word without knowing the 'α' set and 'n'. Value of elements in'α' set is larger random numbers which can't be predicted or calculated by an algorithm and the power over the elements of'α' set makes the encryption more secure.

Further, the cipher value is shared with the third-party server for sorting operation by El-Gamal encryption technique. This encryption technique is used because the parties which know both the secret value of 'n' and 'α' set can decrypt the word of other users so to ensure the security in the computing parties as well we are using El- Gamal encryption technique.

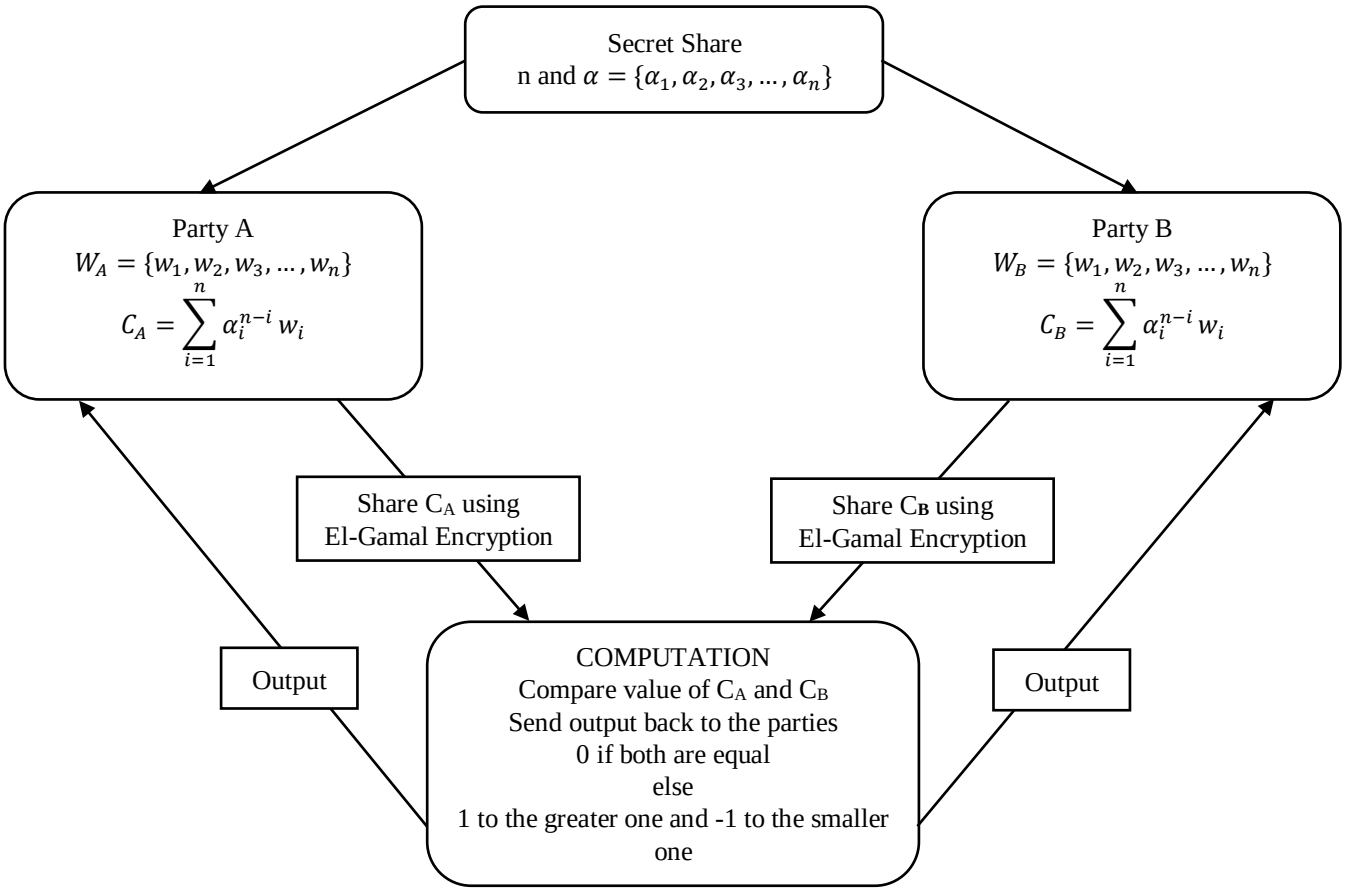

Fig 3: Flow Diagram of Proposed Scheme

B) Time Complexity

The time complexity of proposed scheme is lesser than earlier schemes because it takes the wholeword as an entity. According to the literatureavailable, there is no such algorithm for sorting words. All the available algorithm performs this task only character by character. These approaches first convert each character into ASCII value, so they take high time to sort the word. In the proposed scheme, the computation is performed on the wholeword which makes the system faster and it can sort the enormous number of words easily.

According to literature there is no proper algorithm for word sorting in secure multiparty computation. The proposed secure multiparty computation for sorting words in alphabetic order is less complex and maintain the privacy of data. Use of polynomial function on the whole word makes computation easy, fast and secure. The proposed scheme is not dependent on the length of word or difference in length of the sorting words (i.e. if the first word is of length 'x' and the second word is of length 'y' and x and y are not close to each other).The proposed scheme sorts the word according to the ciphervalue;therefore, it sorts number of words altogether. It does not depend upon the number of input.

REFERENCES

[1] W. Du and M. Atallah, "Secure multi-party computation problems and their applications: a review and open problems," in *Proceedings of the 2001 workshop on New security paradigms*, 2001.

[2] Y. Lindell and B. Pinkas, "Privacy preserving data mining," in *Annual International Cryptology Conference*, 2000.

[3] M. Atallah and W. Du, "Secure multi-party computational geometry," in *Workshop on Algorithms and Data Structures*, 2001.

[4] W. Du and M. Atallah, "Privacy-preserving cooperative scientific computations," in *csfw*, 2001.

[5] G. Aggarwal, N. Mishra and B. Pinkas, "Secure computation of the k th-ranked element," in *International Conference on the Theory and Applications of Cryptographic Techniques*, 2004.

[6] A. Yao, "Protocols for secure computations," in *Foundations of Computer Science, 1982. SFCS'08. 23rd Annual Symposium on*, 1982.

[7] O. Goldreich, S. Micali and A. Wigderson, "How to play any mental game," in *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, 1987.

[8] Y. Gertner, Y. Ishai, E. Kushilevitz and T. Malkin, "Protecting data privacy in private information retrieval schemes," *Journal of Computer and System Sciences*, vol. 60, no. 3, pp. 592-629, 2000.

[9] C. Cachin, S. Micali and M. Stadler, "Computationally private information retrieval with polylogarithmic communication," in *International Conference on the Theory and Applications of Cryptographic Techniques*, 1999.

[10] W. Du, Y. Han and S. Chen, "Privacy-preserving multivariate statistical analysis: Linear regression and classification.," in *Proceedings of the 2004 SIAM international conference on data mining*, 2004.

[11] C. Dwork and K. Nissim, "Privacy-preserving datamining on vertically partitioned databases," in *Annual International Cryptology Conference*, 2004.

[12] A. Shamir, "How to share a secret," in *Communications of the ACM*, 1979.

[13] I. Damgård, M. Fitzi, E. Kiltz, J. Nielsen and T. Toft, "Unconditionally secure constant-rounds multi-party computation for equality, comparison, bits and exponentiation," in *Theory of Cryptography Conference*, 2006.

[14] T. Nishide and K. Ohta, "Multiparty computation for interval, equality, and comparison without bit-decomposition protocol," in *International Workshop on Public Key Cryptography*, 2007.

[15] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," in *IEEE transactions on information theory*, 1985.