Using Gridding Symmetric Encryption for Location Privacy Protection

Ruilin Lai[1], Tao Wang[2], and Yan Zhen Chen[1]
[1]School of Internet Finance and Information Engineering, Guangdong University of Finance, Guangzhou 510520, China
[2]Department of Network Security, Barclays Bank, Stoke-on-Trent ST7 2EP, United Kingdom
Email: 1139656140@qq.com

Abstract —Recently, the major model of location privacy protection is TTP (trusted third party) [1], [2]. However, once the TTP is untrusted or attacked, terminal users have the risk of location privacy leaking. Aiming at the above problems, a GSE scheme (Gridding Symmetric Encryption) comprehensive using gridded coordinate automatic processing and order reserving encryption mechanism is proposed. In gridding operation, the third party does not know exact position of users or destination. In symmetric encrypted operation, the third party also does not know the information of user identity or interest point. In ordered reserving operation, the third party only needs comparison simply. Finally simulation experiment shows the performance of GSE including time cost and communication cost is more efficient.

Index Terms—Cloud computing, gridding system, symmetric encryption, location privacy, order reservation

I. Introduction

With the development of mobile network technology and location technology, LBS(location-based-service) has also been developing rapidly, and attracted more and more attention [3], [4]. Users can get their real-time location coordinates through intelligent mobile terminals with location function, and upload them on related servers, so that they can enjoy all kinds of LBS services provided by LSP (location service providers).

For example, the user can find the best route to somewhere, just entering his current position and destinaiton. Furthermore, he can get to the destination by navigation service in real time.

However, when the server storing user real-time location information is untrusted or attacked, a large number of user privacy location will be leaked [5], [6]. Once these information is used and analysized illegally, the haker can generate users' other privacy, such as working address, living habits etc.

TTP has serious shortcomings as following:

- Once the QS is not trusted or attacked, a large number of user privacy location will be leaked
- QS focuses on privacy information storaging, computing and forwarding, so its tasks become very heavy and it is prone to performance bottleneck [7].
- There is a big security risk in k-anonymous mechanism of TTP.

Aiming at overcoming the above shortpoints, this paper introduces a GSE scheme (gridding sysmetric encrption), including automatic gridding computing, order reserving encrypting, sysmetric encryption based on cryptography.

In the GSE, QS works in close state of data filtering, and it can not get and specific user location information, so QS becomes non-fully trusted. QS simply compares the data in order of query process, so computing overhead become less and bottleneck problem is alleviated in the peak of services. Thus GSE scheme can achieve the security and computing goal.

II. Preliminaries

A. Architecture of GSE

The architecture of the GES contains 3 entity parts: terminal user, QS (query server) and LSP (location service provider) (see Fig. 1).

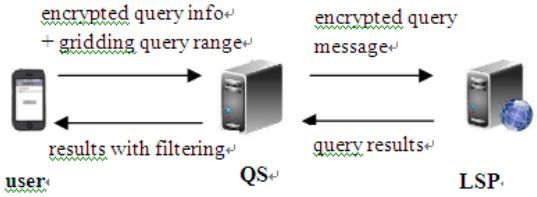

Fig. 1. Architecture of the GSE

User is an intelligent mobile terminal with positioning functions, which can obtain its own real-time coordinates (x_0, y_0). User can set query range, privacy level, gridded anonymous processing and order encryption operation [8], [9].

QS is responsible to pass encrypted query request sent by user to LSP, employed between user and LSP. In addition, QS also stores the gridded query range of order encryption, which is used to filter the query results returned from the LSP, which is the final results back to user.

LSP has a large number of location service databases providing variable location services [10]. When LSP receives request from QS, it will search relative POIs. After processing POIs, LSP return them to QS.

Manuscript received March 23, 2018; revised October 18, 2018.
This work was supported by the Guangdong University of Finance project under Grant No. 16XJ02-07 and name: the research of personal security in mobile network with the example of high school students
Corresponding author email: 1139656140@qq.com
doi:10.12720/jcm.13.11.673-678

III. GSE Scheme Design

There are 5 steps in the GSE scheme including grid process of user, storage and forwarding of QS, routes searching of LPS, index table building of QS, and tracking movement of user (see Fig. 2).

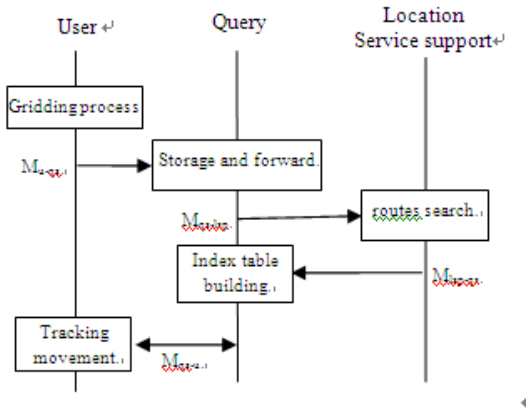

Fig. 2. Five steps of GSE

The following Table I shows symbols and related description of GSE [11], [12].

TABLE I. SYMBOLS AND RELATED DESCRIPTION OF GSE

Id	Symbol	Description
1	n	degree of navigational accuracy
2	p	degree of privacy protection
3	PO_s	Start position of user
4	PO_e	destination of user
5	Pri_{lsp}	Private key in LSP
6	Pub_{lsp}	public key generated by SP
7	key_{op}	Private key of ordered reserving encryption
8	$key_{u\text{-}lsp}$	Symmetric key generated by user
9	(x,y)	coordinate
10	S	gridding area
11	time	time stamp between user and QS
12	r	route as <start point, destination>
13	r’	r encrypted by ordered reservation
14	sset	Set of expanded start point
15	dset	Set of expanded end point
16	R	Set of all routes

A. Gridding Process of User

At the beginning, user should set up his destination position (x,y), privacy protection degree p and navigational accuracy degree n [13]. For example, when a user whose position is (x_s, y_s), wants to navigate to destination (x_e, y_e), within privacy degree 0 and accuracy degree 8, he should input message as following expression.

$$\{start=(x_s, y_s), destination=(x_e, y_e), p=0, n=8\}. \quad (1)$$

Privacy degree is used to protect user real points in LSP, which has two values 0 or 1. 0 means result route is direct line and 1 means result route is undirested lines. Obviously undierscted line has higher private security as LSP cannot know user’s start point and destination [14], [15].

Then, based on the user's current position (x_s, y_s) and his destination (x_e, y_e), the surrounding area is divided into a square grid S as following expression. Coordinates (x, y) are transferred to gridding coordinates points PO_s and PO_e.

$$S \leftarrow \{(x_s, y_s), (x_e, y_e), n\} \quad (2)$$

For example, in gridding coordinate system, we can see query range S is {(x_s, y_s), (x_e,y_e), 8}, 8 of which is the number of grids in row or in column, and user gridding points are PO_s = (2, 5) and PO_e = (6, 2) (see Fig. 3).

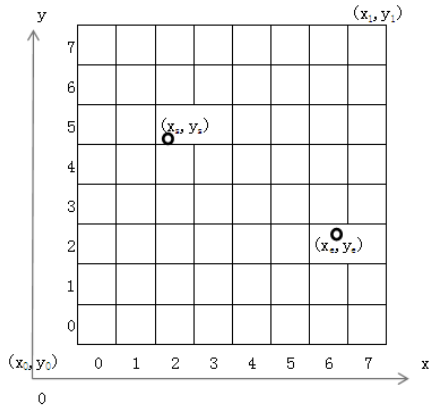

Fig. 3. Gridding coordinate system of start point and destination

Secondly, in gridding system, above 2 points are encrypted and put into set pset, by order reserving encryption algorithm key_{op} as the following expression.

$$pset \leftarrow \{key_{op}(PO_s), key_{op}(PO_e)\} \quad (3)$$

Thirdly, in order that user real position cannot be found in query message, a method of extending points set is used.

Distance parameter DIS is generated by random in certain range. And then original 2 points (PO_s and PO_e) duplicate and move a distance of DIS by random direction to new location points (PO_{s1} and PO_{e1}) [16], [17].

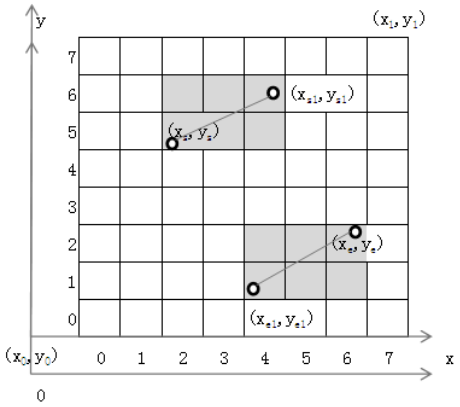

Fig. 4. Sets of points expending

For example, as shown in Fig. 4, the new additional start point is PO_{s1} (4,6) and new additional destination is PO_{e1} (4,1). The 2 relative shadow parts are made by 2 diagonals, and the 2 endpoints of each diagonal are start points or destinations. Thus the shadow parts contain new start point set sset and new destination set dset respectively (see Fig.4).

The following expression is shown as expanded sets rebuilt.

$$\text{sset} \leftarrow \{\text{POs}, \text{POs1}\} \quad (4)$$

$$\text{dset} \leftarrow \{\text{POe}, \text{POe1}\}$$

Fourthly, Aiming at data transmission security, symmetric key of this query key_{u_lsp} is generated randomly by DES encryption method, which will be used in transmission of specific results between user and QS in final step. In order to prevent verify data integrity, a time stamp is also generated.

A public key pub_{sp} from LSP is given to user, which is generated by RSA key exchange protocol. Parameter Pub is made by the encryption query set. Thus, a query message into M_{u_qs}, is combined with some parameters and sent to QS as following expression.

$$\text{Pub} \leftarrow pub_{sp}(\{S, \text{sset}, \text{dset}, key_{op}, 0\}) \quad (5)$$

$$M_{u\_qs} \leftarrow \{\text{Pub}, \text{pset}, key_{u\_sp}, \text{time}\}$$

B. Storage and Forward of QS

When QS receives the query message M_{u_qs}, it keeps the information pset, time, and key_{u_sp}, which are used to compare cooperates, validate transmission time, and symmetrical encrypt results in final step. Then QS rebuild a new query message M_{qs_lsp}, to LSP, which is shown as following expression.

$$M_{u\_lsq} \leftarrow pub_{sp}(\{S, \text{sset}, \text{dset}, key_{op}\}) \quad (6)$$

C. Routes Searching of LSP

When LSP receives query message M_{u_lqs}, a private key pri_{sp} is used to decrypt the above expression 8. The following parameters can be generated: grid surrounding area S, expanded start points set, expanded destination set, ordered reservation encryption key_{op}, and privacy degree.

After building the expanded point sets on gridding S, LSP looks up all routes of points in coordinate database, matching the best directed routes or undirected routes. Every route R_i has 3 critical information r_i, r_i' and w_i, which means that r_i is endpoints as identical key, r_i' is ordered reservation encryption points, and w_i is its context.

The line symbols represent result routes, including route R_1, R_2, R_3, and R (see Fig. 5).

For example, in the above figure, positions of route 1 is <(2,5), (6,2)>, route 2 is<(2,5), (4,1)>, route 3 is <(4,6), (6,2)>, and <(4, 6), (4,1)>. These route position coordinates are input into r_i (i=1,2,3,4), and then r_i' is generated by key_{op} with r_i. Relative detailed description are also put into w_i. Finally, <r_i', w_i> are put into message R_i. The following expression9 is shown as following.

$$r_i \leftarrow <(x_{s\_i}, y_{s\_i}), (x_{e\_i}, y_{e\_i})> \quad (7)$$

$$r_i' \leftarrow key_{op}(r_i)$$

$$R_i \leftarrow < key_{op}(w_i), r_i'> \quad (1<=i<=t)$$

(when R is directed route, t is 4. when R is undirected route, t is 8)

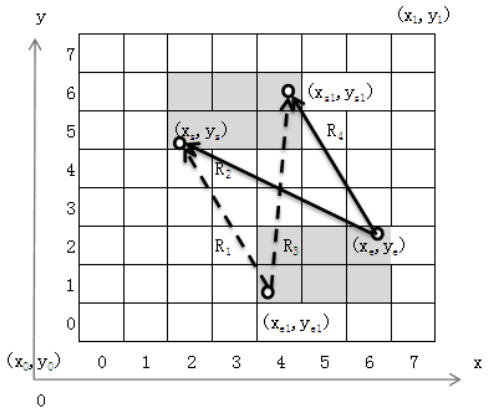

Fig. 5. Query results of routes

When all 4 routes are finished, they will be put into message M_{lsp_qs} and returned to QS as following expression.

$$M_{lsp\_qs} \leftarrow \{R_i\} \ (1<=i<=t) \quad (8)$$

D. Index Table building of QS

When QS receives the return message M_{lsp_qs}, each r_i' (1<=i<=t) of routes is compared with the earlier stored point pset in ordered reserving encryption state, and only one demanded route R_2 <(x_s, y_s), (x_e, y_e)> are kept while others deleted. In gridding system, line R_2 is converted to points set R_2' composed of lots of points (see Fig. 6).

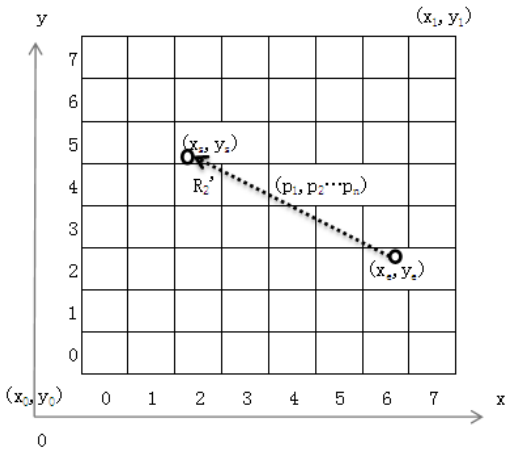

Fig. 6. R2' composed of lots of points

Parameter w_2 of R_2 is converted to routing index table combination with ordered encryption and symmetric encryption. In routing index table, each element has 2

parameters key_{op} (p_j) which is location of each point encrypted by key_{op}, and $key_{u\_lsp}(des_j)$ which is the description of relative point encrypted by key_{u_lsp}, including movement from previous point, traffic condition, nearby buildings etc as following expression.

$$\text{Index table} \leftarrow < key_{op}(p_j), key_{u\_lsp}(des_j) > (1<=j<=n) \quad (9)$$

E. Tracking Movement of User

After QS builds routing index table, user can communicate with it for query of next movement.

In QS, pset from user is compared with each p_i by judging the relative size of numbers in the close state based on routing index table. Due to ordered reservation encryption, it is easy to pick up the current point and the next movement. And then, the next movement $key_{u\_lsp}(des_i)$ is sent back to user. After decrypted by symmetric decryption key_{u_lsp}, user obtains the information of next movement.

Only when user's current position is not in R_2, or timestamp is over, the user needs to send M_{u_qs} to LSP for requesting a new route again, processing the first step again as a loop.

IV. Performance Testing

Experiment platform is a 64 bit PC of Win10 operation system, with a CPU of 2.1 GHz and a memory of 4 GB. Software platform is SQL Server 2012 Business Intelligence database server, and Visual Studio Professional 2012 as C# development tools. During experiment, the size of map is 50000 x 50000, with 1 meter as a unit. The current position of user and destination is generated based on Students Behavior database. The third part (TTP) is simulated by inter-process communication of PC[18], [19].

A. Test for Changes of Distance of 2 Points

GSE performance includes time cost and communication cost [20], and these performance results are focused by a series of 2 points distance changes and the number of grid partition number n changes, based on the directed route setting. From the following Fig. 7 and Fig. 8, the more parameter distance or parameter n is set, the more time cost and communication of GSE become.

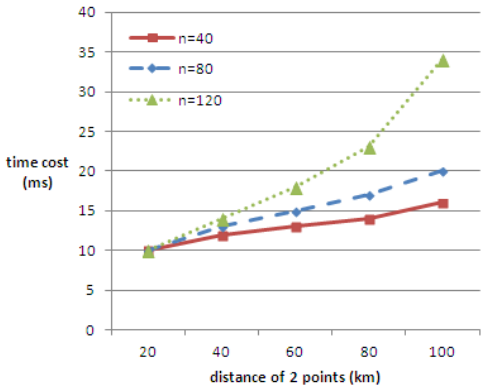

Fig. 7. Time cost changes associated with different distances in different grid partition

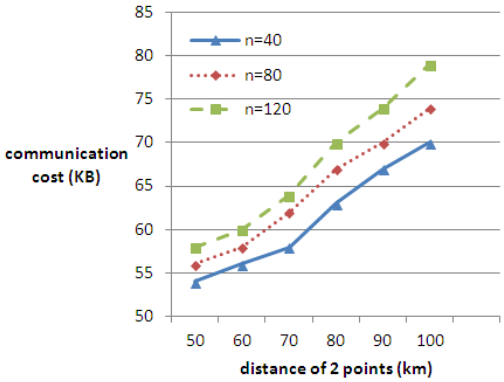

Fig. 8. Communication cost changes associated with different distance in different grid partition

The reason is that when distance is set larger, query route becomes larger, which leads to time cost and communication cost becoming greater. When grid partition n is set larger, the number of overall grid becomes more and the number of processing grid becomes more, which leads to time cost and communication cost becoming greater.

In Fig. 7 of time cost, the time cost increases greatly, because the overall time cost of the system is mainly affected by 3 entities calculation including user, QS and LSP. With the relevant parameters setting larger, even thought the volume of user and QS calculation has less grew, the number of processing grid and processing POIs are increasing greatly. It leads to more encrypted operations will be processed in LSP. It is the reason that the calculation of LSP increases obviously and overall time cost grows significantly.

B. Test for Changes of Directed Route or Undirected Route

When the privacy level is set 0 or 1, it means routes generated in LSP are directed or undirected. The goal of this setting is that a number of routes are generated one of which is just user's demand, and LSP cannot which one is user's real route.

This experience focuses on the relationship between time cost or communication cost and distances changes, in category directed route and undirected route, based on n-80.

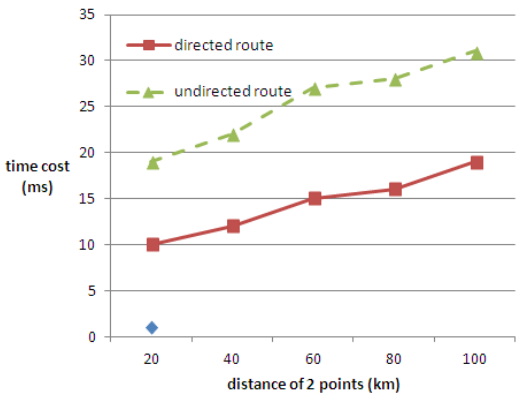

Fig. 9. Time cost changes associated with different distances in directed route or undirected route

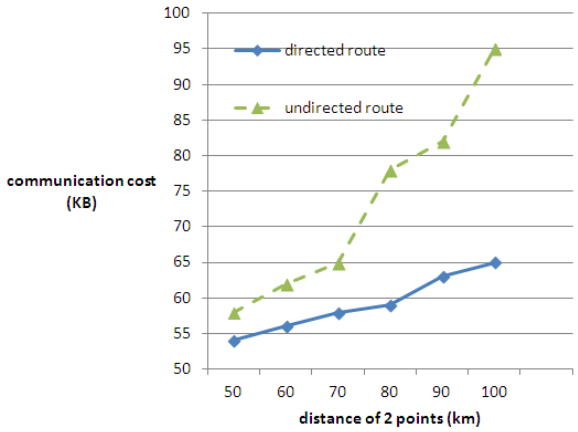

Fig. 10. Communication cost changes associated with different distance in directed route or undirected route

From Fig. 9 and Fig. 10, it shows that the increase of the time cost is significantly higher than the communication cost. When distance becomes larger and the directed route is set, the volume of return routes the processed will become more significantly.

In Fig. 9 of communication cost, even thought relevant parameters are set larger, the number of storing routes and transmission routes is not change, so as not to impact the space overhead. So it is reason that communication cost does not grow obviously.

V. Conclusions

The problem of privacy protection based on location services is more and more popular recently [21]. Because of the shortcomings of TTP, this paper proposes a GSE scheme, comprehensive using gridded coordinate automatic processing and order reserving encryption mechanism. The GSE converts specific coordinates to grid model, which is used for process of close state in the entire query duration, in order to avoid extreme situations in expanded query, and improve the degree of privacy protection. What is more, because the QS only needs to compare 2 coordinate size of close, it will reduce the calculation pressure in the peak of services efficiently. So GSE is more efficient and secure solution of location privacy protection, which can fix the non-trusted TTP problem.

Acknowledgment

This work is sponsored by Guangdong University of Finance (project number: 16XJ02-07, name: the research of personal security in mobile network with the example of high school students).

References

[1] J. Li, *et al*., "Privacy-preserving data utilization in hybrid clouds," *Future Gener Comput. Syst*., vol. 30, no. 1, pp. 98–1066, 2014.

[2] Y. Li, *et al*., "Privacy preserving cloud data auditing with efficient key update," *Future Gener Comput. Syst*., vol. 78, pp. 789–798, 2016.

[3] Y. Wang, "Privacy-preserving data storage in cloud using array BP-XOR codes," *IEEE Trans Cloud Comput*., vol. 3, no. 4, pp. 425–436, 2015.

[4] A. Boldyreva, N. Chenette, Y. Lee, *et al*., "Order-preserving symmetric encryption," in *Proc. Annual International Conference on the Theory and Applications of Cryptographic Techniques*, 2009, pp. 224-241.

[5] F. Kohlmayer, *et al*., "A flexible approach to distributed data anonymization," *J Biomed Inform*, vol. 50, pp. 62–76, 2014.

[6] S. Goryczka, *et al*., "m-Privacy for collaborative data publishing," *IEEE Trans Knowl Data Eng*., vol. 26, no. 10, 2014.

[7] G. Zhang, *et al*., "A historical probability based noise generation strategy for privacy protection in cloud computing," *J Comput. Syst. Sci*., vol. 78, pp. 1374–1381, 2012.

[8] J. V. Nayahi and V. Kavitha, "Privacy and utility preserving data clustering for data anonymization and distribution on Hadoop," *Future Gener Comput Syst.*, 2016.

[9] J. V. Nayahi and V. Kavitha, "An efficient clustering for anonymizing data and protecting sensitive labels," *Int. J Uncert Fuzziness Knowl Based Syst*., vol. 23, pp. 685–714, 2015.

[10] S. Nan, J. Chun-fu, *et al*., "Approach of location privacy protection based on order preservation encryption of grid," *Journal on Communications*, vol. 7, pp. 78-88, 2017.

[11] M. Ahmadian, A. Paya, and D. C. Marinescu, "Security of applications involving multiple organizations and order preserving encryption in hybrid cloud environments," in *Proc. IEEE International Conference on Parallel & Distributed Processing Symposium Workshops*, 2014, pp. 894-903.

[12] T. S. Hsu, C. J. Liau, and D. W. Wang, "Logic, probability, and privacy: A framework," in *Turing-100*, 2012, pp 157–167.

[13] Q. Zheng, *et al*., "VABKS: Verifiable attribute-based keyword search over outsourced encrypted data," in *Proc. IEEE Conference Computer Communication*, 2014.

[14] D. Rebollo-Monedero, *et al*., "From t-closeness-like privacy to postrandomization via information theory," *IEEE Trans Knowl Data Eng*., vol. 22, pp. 1623–1636, 2010.

[15] C. Phua, V. Lee, K. Smith, *et a1*., "A comprehensive survey of data mining-based fraud detection research," *Artificial Intellgence Review*, 2010, pp. 1-14

[16] B. Waters, *et al*., "Conjunctive, subset, and range queries on encrypted data," *Theory Cryptogr*, vol. 4392, pp. 535–545, 2007.

[17] R. Agrawal, J. Kiernan, R. Srikant, *et al*., "Order preserving encryption for numeric data," in *Proc. International Conference on Management of Data*, 2004, pp. 563-574.

[18] F. Amiri, *et al*., "Hierarchical anonymization algorithms against background knowledge attack in data releasing," *Knowl-Based Syst.*, vol. 101, pp. 71–89, 2015.
[19] C. Gentry and S. Halevi, "Hierarchical identity based encryption with polynomially many levels," in *Proc. Theory of Cryptography Conference*, 2009, pp. 437-456.
[20] X. Zhang, *et al*., "Proximity-aware local recoding anonymization with mapreduce for scalable big data privacy preservation in cloud," *IEEE Trans Computing*, vol. 64, no. 8, pp. 2293–2307, 2015.
[21] L. Wen-Yang, *et al*., "Privacy preserving data anonymization of spontaneous ADE reporting system dataset," *BMC Med Inform Decis Mak.*, 2015.

Ruilin Lai was born in Guangdong Province, China, in 1981. He received the B.S.c degree from the University of Ulster (United Kingdom) in 2005, and the M.S.c degree of Telecommunication and Internet Systems from the University of Ulster (United Kingdom) in 2007. He worked at Invision Software Ltd in UK as software engineer in the past. He is currently pursuing the search of cloud computing and network security.