Bartering Method for Improving Privacy of LBS

Abdullah M Basahel[1], Adnan Ahmed Abi Sen[2], Mohammad Yamin[3], Sarah Alqahtani[4]

[1]Department of MIS, Faculty of Economics and Administration, King Abdulaziz University, Jeddah, Saudi Arabia
[2]College of Computer and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
[3]Department of MIS, Faculty of Economics and Administration, King Abdulaziz University, Jeddah, Saudi Arabia
[4]Independent Researcher, Saudi Arabia

Summary
Privacy issues are considered to be the greatest challenge facing the future of new technologies. For this, much research has been presented in this domain and researchers have proposed many approaches and methods for protecting privacy and security, but there are many open problems in these approaches related to efficiency and performance. This research provides a review for the previous main approaches, and then presents a new method: "Bartering Technique." This is an enhancement for cooperation, caching, and dummies techniques, where Bartering Technique increases the level of privacy and decreases the cost and performance compared to the previous approaches. Through simulation and comparison, we proved the superiority of Bartering Technique on traditional dummies and enhanced-caching techniques.

Key words:
Privacy, Security, LBS, POI, Bartering

1. Introduction

Most recent mobile or web applications in addition to new systems depend on location-based services (LBS), one of the most important apps being Point of Interest (POI), where users can search for the nearest place for their current location (e.g. hospitals, restaurants, hotels, schools, etc.). That can be achieved by general structure of LBS and general scenarios [1].

First, the user's mobile device determines its coordinates depending on Global Positions System (GPS). Second, the application on the device sends the user's query to the LBS server provider (SP). Third, SP has a database of all places and their locations, like Google Map, so SP will find all results that meet the user's query (the places of same wanted type, and how far from the user's position less than selected range). Fourth, SP returns the results to the user [1-2].

So, location-based services (LBS) has become a very important part in many of the recent smart applications which enable them to provide a new dimension of services and a higher level of flexibility for using these applications by users searching for points of interest (POI), delivering and dropping requests, contacting, navigation, managing traffic, and monitoring patients, employees, children, etc. [3, 4]

Manuscript received February 5, 2019
Manuscript revised February 20, 2019

However, on the other side, this type of service makes the users' data vulnerable to infiltration, and their sensitive information can be disclosed. For example, if an attacker knows information about where you are and when, that means your private information can be revealed like what your job is, your social status, religion, friends, home location, medical information or records, and so on [5-6].

Since the query in LBS applications contains five main parts (ID, Timestamp, Current Location, Query's Type (POI), and Range), see Figure 1, each privacy approach attempts to protect one or more of these parts according to the type of application [7].

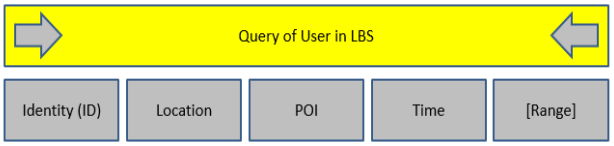

Fig. 1 Parts of query in LBS.

There are differences between security and privacy concepts as Figure 2, Table 1 & 2 depict [5-7].

"**Security**" has three main attributes which are: Confidentiality, Integrity, and Availability.

Meanwhile, "**Privacy**" has three different attributes which are: Traceability, Link-ability, and Identifiability.

Table 1: Security Terms

Security Term	***Description***
Confidentiality	A set of rules or a promise that limits access or places restrictions on certain types of information such as using encryption to prevent showing data from any unauthorized users
Integrity	Maintaining and assuring the accuracy and consistency of data over its entire life-cycle. That means preventing and detecting any attempt to edit data
Availability	The system is available 24/7

Table 2: Privacy Terms

Privacy Term	***Description***
Traceability	The attacker will be able to record the accurate location of a user by the time
Likeability	The attacker will be able to link specific data to a specific user
Identifiability	The attacker can sufficiently identify the entities within the system

Fig. 2 Security Triads V.S. Privacy

However, the most important difference between both concepts is that in the security, a user will be trusted with SP, but in privacy that isn't necessary, where SP can pose the greatest threats to the user's privacy because it has all the data about this user. So, this research is focused more on privacy than security, and especially on the previous approaches and methods that tried to address this issue.

Privacy is one of the greatest topics of interest in recent times, which is frequently discussed with regard to new applications and technologies. It refers to the right of users to determine who can see or use their data, and when that will be, in addition to where it will be used, and why [7].

This research focuses on Privacy and its contributions are:

- Present an overview about the privacy issue, and its approaches, especially in POI applications which are considered the significant type for LBS.
- Explain the difference between privacy and security concepts and review the popular kinds of attacks on privacy.
- Provide a new technique for preserving privacy in POI applications, which achieves a greater level of privacy and reduces the cost and headache for all parties.

2. Literature Review

Much research and many techniques are provided for preserving privacy of LBS and its applications. The authors in [8] showed that such information on these sensitive aspects (such as user location, or POI) could be obtained by an attacker, where they could track the locations of users or analyze their queries and their POI. After gathering sensitive data about the victim, the attacker can detect a lot of important information about the user such as when they are in or out of their home, where they are now, what is their job, if they are rich or poor, their interests, their religion, their ethics, etc., and some authors point out far more than that. For that, many researches have concerns about protecting security and privacy of data in like of these applications (POI) that rely on LBS [8].

Many of the approaches and methods are provided for preserving privacy in LBS and preventing attacks. A classification for the proposed approaches is provided in [8-9]. The authors here classified the approaches to three groups: one has the approaches that depend on another server to protect a user's privacy, another has the approaches that depend on the users themselves, and the last has the approach that relies on other users or cooperation between many of the users. In the following Figure 3, we have classified the proposed approaches in POI used in research.

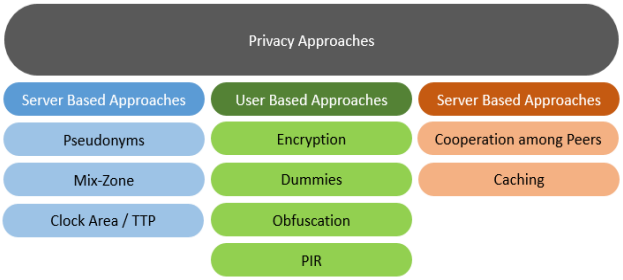

Fig. 3 Classification of Privacy Approaches

Table 3: Privacy Approaches

Approach	***Description***
Encryption [10]	This approach considered the server provider (SP) is trusted and protects the privacy of data from outside attackers by using symmetric or asymmetric Encryption. This technique is very strong against outside attackers, but SP itself which can be an attacker also, and can analyze the collected data of users to detect sensitive information about them. But this technique can be problematic for the resources of user's device and on SP itself
Pseudonyms & Mix Zone [11]	There is also a simple technique where the user can change his ID and use a nickname when he is dealing with SP, but this approach isn't effective if the attacker has the ability to collect many queries about the user, because this attacker will analyze this data to detect important information about user which can help to disclose his identity. An enhancement for this technique is Mix-Zone where the user has to change his nickname in each area and doesn't use only one pseudonym.
Third trusted party (TTP) & Cloak Area	This approach has many techniques, all of which depend on other servers to protect user privacy from SP. The first technique is Anonymity for users' ID where TTP will send the query to server instead of the user. In this case SP can't collect any data about this user [12]. Second technique is clacking area of users' locations [13]. Here the city will be divided into many cells (Clack areas), and in each cell there is an "Anonymizer" which collects the queries of all users in its area and then sends them as a set in the same time (as one query) to SP to prevent them to link any information with any user. Moreover, the TTP will send only the coordinates of Clack area, not the accurate locations of each user. But the user has to trust in TTP also, which means he only shifted the problem from SP to TTP [14].

Dummy Approach [15]	This approach relied on the user himself by generating many dummies besides his real query, and then sending them as a set together to SP which will be not able to determine which the real one from this group of queries is. Note, all queries will have the same location and different type (Query-Dummy), or same type and different locations (Locations-Dummy). But the process of generating good dummies that prevent SP to distinguish which is the real one form others, is still an open problem in this domain.
Obfuscation [16]	Here, the user will also not send his accurate location to SP with his query. Instead of that, the user will add noise to his location and then send it to SP. In this case SP can't track the user location. But that surely will effect the accuracy of results, SP can draw a path for users and reveal their accurate locations.
Private Information Retrieval (PIR) [17]	Here the user will retrieve a huge part of data from SP to prevent it to know which part that user needs specifically, for example, all data about hospitals in the selected city, so SP can't know any information about user location. Moreover, this approach can employ encryption and divide the area for cells with numbers, and the user can request many of these numbers in one query. But this approach is very heavy and can be problematic and cause an overload for the user, network, and SP.
Cooperation between users/ peers [18]	User collaborates with other users to protect their privacy from SP itself. Where each one of them can save their query results in their cache, and if any other user asks them about similar queries they can answer that directly instead of communicating with SP. Other methods in this approach depend on sharing all users in a specific area for the same location and then sending it to SP to make it unable to distinguish between peers. But the trust and dealing among peers is still an open problem in this domain.
Hybrid techniques [18]	Integrating between many techniques or using caching with one or more of the previous methods. But the position of the cache in the system, and enhancing the cache hit-ration is still an open problem in this domain.

Table 3 provides a review of these approaches one by one and explains the advantages and disadvantages of each.
Briefly, many approaches and methods are provided for preserving privacy of LBS and POI apps, some of them dealing with outside attackers, and another considered SP itself as an attacker. However, there isn't a perfect technique that can provide comprehensive protection for user's privacy. In the next part we will try to propose a new technique which will address some drawbacks of previous techniques and enhancing the level of privacy of users.

3. Proposed Technique (Bartering technology)

The idea came as a development of Hiding in the Crowd Approach, but in a new way that addresses the previous problems in both Hiding in the Crowd and the Dummies approach. Figure 4 depicts the main steps in the technique.

3.1 Drawbacks of Previous Techniques

The dummies approach is one of the most famous privacy protection techniques used in location-based services, especially when looking for points of interest (POI). The user sends a set of queries (K) to the SP instead of sending one query. But unfortunately, this approach has a lot of drawbacks as well [19-20]:

- Increasing the value of K means increasing the load on the user, the server, and the network because of sending multiple numbers of queries from sending a single query. In other words, increased privacy is proportional to the increased cost and performance in this approach.
- Dummies algorithm (or user-generated fake queries) is a cumbersome and incomprehensible process because of the ability of the server or external attacker to monitor frequent queries from the user (continuous | consecutive) and thus their ability to detect many of the phantom queries, especially if the attacker has sufficient knowledge of the map of the area in which the user is moving, where some sites that are not logical (e.g. over a sea, a mountain area, etc.) can be deleted.

As a development of this approach, another approach (Hiding in the Crowd) has been presented: it is using cache on each user's device to temporarily store its query results. And then, rely on a collaborative approach among users to reduce the first problem in the dummies approach, which concerns overload [19].
The main idea of the previous approach is that the SP is the greatest threat to the privacy of the user if the provider is malicious, as it collects all the data and user queries, thus reducing the number of connections in the service provider means improving performance and reducing cost while increasing the level of user privacy.
The assumption is that users in a particular area are often looking for similar things. This has been proven by analyzing real data in Google Maps; so, user A must first send a query to another user in their area before sending it directly to SP with a group of dummies. If B has the answer to that query within its cache (i.e. it has already been asked on the same subject), it will send the answer directly to user A and therefore will not need to contact the service provider in this case [19-20].

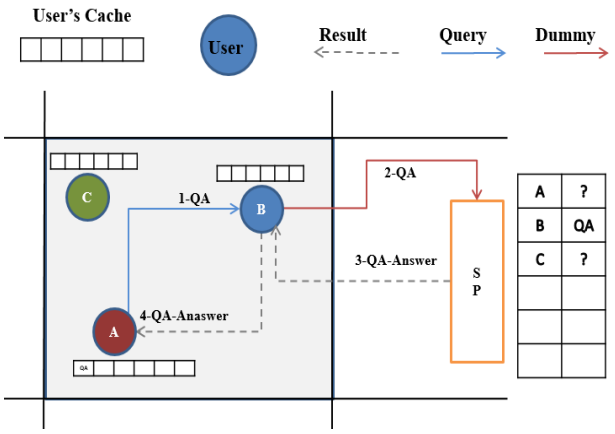

Fig. 4 Main Steps of Bartering Technique

Thus, the number of connections to the service provider will be reduced and the amount of information generated on user A will be reduced, which means an improvement in privacy and performance at the same time. Although, Hiding in Crowd has improved the level of privacy and performance, but only in cases where it assumes that another user in the same region has an answer to the same query. Otherwise, the first dummies approach will be used again when communicating with the SP.

3.2 Steps of Barter Technique

The idea is summarized as follows:
First, user A selects one of the users around them randomly who use the same proposed technique, and let it be B.
Second, user A sends their query AQ to B where the query will be a point of interest and user A's location.

A. Send (AQ, B)

Third, user A Stores User B in their memory with an initial score of 1 indicating the number of times of collaboration.
Fourth, B receives the query sent from A.
Fifth, B looks for an answer to AQ in their Cache.
Sixth, if B finds an answer to AQ, B will send the answer to A and the process will end here without having to contact the SP.

B. Send (B_ Result, A)

Seventh, if B does not find the result of AQ, B will send the query to SP on behalf of A.

B. Send (AQ, SP)

Eighth, the SP will receive one query coming from B, the SP will answer it and return the result to B.

SP. Send (Result, B)

Ninth, the SP will store false information about B in its database because the query does not belong to B and can therefore be considered as false (AQ will be as dummy for B); i.e. B misleads the SP and gives it false information.

Tenth, B will send the result to A which has protected its privacy completely from the SP.

B. Send (Result, A)

Eleventh, A can change its alias to a new name to increase the level of privacy from dealing with neighbors in subsequent times because the normal user can't track the different names of another user as the provider does.
This process is repeated in the case of future queries, but not to deal with the same user B for more than a certain level of queries, so that B does not pose a threat to user privacy either. Do not forget that A may be B in other cases.

3.3 Advantages of Barter Technique

- The proposed approach contributes to a lot of load on the user, the SP and the network so that only one query will always be sent over the network.
- The proposed approach obscures direct communication between the user and the SP, unless sending a query to another user, which means complete protection of the user from the SP and giving them completely false information.
- B will send a variety of completely random queries to the SP because it belongs to different users and thus will increase the level of privacy more and more without additional impact on performance.
- This approach achieves all the benefits of the Hiding in the Crowd approach and avoids its negatives as well.
- The approach provides a solution to the problem of having a smart algorithm to generate dummies since each query for another user is considered as dummy for the sending user.
- Achieves a common benefit for each of the collaborating users.

3.4 Disadvantages of Barter Technique

- The process of communication and reliability of another user, although the degree of risk does not compare with the degree of risk of communication with SP, but that does not mean absolute protection of the user.
- The matter of the availability of other users and their acceptance of the process of cooperation regardless of their usefulness.
- The waiting time cannot be guessed by another user, but a certain time can be specified if it is possible to resend the query to a new neighboring user.
- The question of user B returning User A incorrect results can't be adjusted, although they are rare but can occur, some of the solutions that can be used here are the issue of reputation.

4. Attacks that could be addressed by the proposed approach

Semantic Context: In the event that the attacker has personal information about the user such as their age or job, the attacker may use this information to reveal the user's query. In the proposed technology, the user does not deal directly with the server in the case of sending queries for another user. Even if the attacker reveals that these queries do not belong to the sender, this will not affect the privacy of the user [21-22].

Path Tracking: It can be useful in Traditional Dummy but in the proposed system, the user each time sends one query to a different user, so the queries will be very random and non-threaded and thus untraceable.

Historical Data: The attacker here gathers a lot of data or queries for the user over a period of time and then analyzes and discovers new information about the user. But in the proposed technology, all the information stored on the user will be information that does not belong to them and therefore the results of the analysis will be incorrect and will not affect the user's privacy.

Inversion Attack: Here the attacker has knowledge of the algorithm used for protection, but in the proposed technique, even if the attacker knew that the next query from a particular user does not belong to him, but is for another user, they will not be able to know this user at all and will not affect the privacy of one of the users.

Some notes to be taken into account in privacy protection [21-22]:

Diversity: The areas which are queried from, should not be one of a single interest. For example, all POI are related to the medical side and thus it is easy for the attacker to guess which query is the real one and which query is false. However, in the proposed technology, this issue was bypassed by the user sending only real queries to the SP, which are diverse but not specific to the user, thus increasing the level of privacy.

Congestion: The number of users in a cell must be relatively large to ensure greater protection and greater diversity.

Knowledge about the map: This observation is in order to generate dummies in real areas and not generate them in areas that are illogical and thus easy to filter by the attacker, but this observation was exceeded in the proposed technique because all queries are real queries and therefore are logical queries.

5. Results and analysis

There are several key factors to compare two technologies or approaches in the field of privacy protection in LBS, particularly those related to POIs [18-22]

A. The Coefficient of Privacy

This is determined by the amount of correct information generated by the attacker (which can be the same SP) for each user and this parameter can be determined by two basic parameters:

K-Anonymity is the number of real queries compared to the phantom queries coming to the server.

$$\text{K-Anonymity} = 1/K \quad (1)$$

K number of queries that send to SP in each send operation

Entropy: This is a standard to the extent that the server confirms that this information is related to the user or not and is linked to the possibilities. It can be limited between 0 and 1. In privacy, the higher the value, the higher the degree of privacy.

$$E=\sum pi*\log2\ (pi) \quad (2)$$

Pi is the probability of Qi is related to user

Other criteria such as **Ubiquity** and the **Estimation Error** are also related to this parameter, but all depend on the entropy concept.

B. The Coefficient of Performance and Cost

Determined by several metrics:

Number of queries sent to the service provider in each dispatch or request.

Cache Hit Ratio, which means reducing the number of contacts with the SP.

Amount of information which is required to be returned by the SP.

Accuracy of the results and the amount of required treatment.

Here is a simple comparison between our proposed technology and the traditional dummies approach, and then a comparison between the proposed technology and the Enhanced Cache approach that uses cache with dummies to reduce the number of connections with the server [19-20].

It has been assumed that the hit rate is 35% and K = 5 and that each user sent once

Figure 5, and 6 confirm the superiority of the proposed technique compared to both of the previous two techniques according to privacy and performance levels, which also is logically confirmed by the scenario and interpretation discussed in the previous sections.

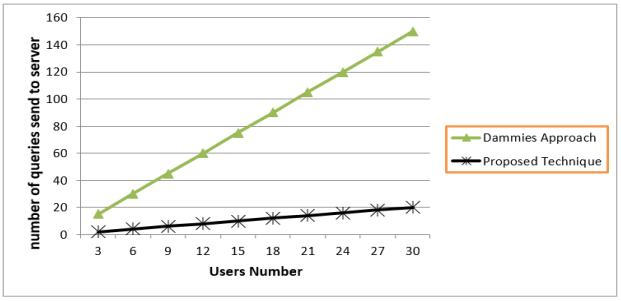

Fig. 5 the number of sent queries compared to the number of users

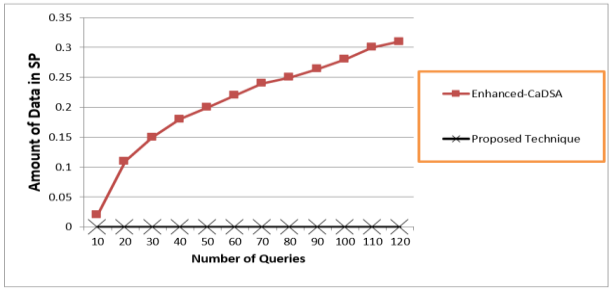

Fig. 6 the amount of correct information generated by the service provider with increases the number of queries.

Summary

This research provides a new technology to protect the privacy of the services based on the location, especially in the systems of POIs, through the barter between users to form dummies in a new way, in addition to using a cache for each user. The proposed technology provides effective solutions to the problems of the dummy approach which related to the dummies algorithm as well as to the reduction of load on the user, the server and the network. Comparison with previous approaches has demonstrated the superiority of the proposed approach in terms of level of privacy as well as level of performance and cost.

References

[1] Mohapatra, D., & Suma, S. B. (2005, January). Survey of location based wireless services. In Personal Wireless Communications, 2005. ICPWC 2005. 2005 IEEE International Conference on (pp. 358-362). IEEE.

[2] Phadnis M, Kadam GV (2016) Efficient geosocial application query processing with privacy preserving policy. Int J Eng Dev Res 188–194

[3] Yamin, M., Basahel, A. A., & Abi Sen, A. A. (2018). Managing Crowds with Wireless and Mobile Technologies. Wireless Communications and Mobile Computing, 2018.

[4] Fouz, F., & Sen, A. A. (2016). Performance and Scheduling Of Hpc Applications In Cloud. Journal of Theoretical & Applied Information Technology, 85(3).

[5] Abomhara M, Køien GM (2014) Security and privacy in the internet of things: current status and open issues. In: Privacy and security in mobile systems (PRISMS), 2014 international conference. IEEE, pp 1–8

[6] Al-Rahal, M. S., ABI SEN, A. D. N. A. N., & Basuhil, A. A. (2016). HIGH LEVEL SECURITY BASED STEGANORAPHY IN IMAGE AND AUDIO FILES. Journal of Theoretical & Applied Information Technology, 87(1).

[7] bi Sen A, Albouraey F, Jambi KA (2017) Preserving privacy of smart cities based on the fog computing. In: Smart societies infrastructure, technologies, and applications (SCITA), Springer

[8] Sen, A. A. A., Eassa, F. A., Jambi, K., & Yamin, M. (2018). Preserving privacy in internet of things: a survey. International Journal of Information Technology, 10(2), 189-200.

[9] Wernke M, Skvortsov P, Du¨ rr F, Rothermel K (2014) A classification of location privacy attacks and approaches. Pers Ubiquit Comput 18(1):163–175

[10] Solanas, A., & Martínez-Ballesté, A. (2007, June). Privacy protection in location-based services through a public-key privacy homomorphism. In European Public Key Infrastructure Workshop (pp. 362-368). Springer, Berlin, Heidelberg.

[11] Liu, X., Zhao, H., Pan, M., Yue, H., Li, X., & Fang, Y. (2012, March). Traffic-aware multiple mix zone placement for protecting location privacy. In INFOCOM, 2012 Proceedings IEEE (pp. 972-980). IEEE.

[12] Shokri R, Troncoso C, Diaz C, Freudiger J, Hubaux JP (2010) Unraveling an old cloak: k-anonymity for location privacy. In: Proceedings of the 9th annual ACM workshop on privacy in the electronic society. ACM, pp 115–118

[13] Jagwani P, Kaushik S (2016) Secure cloaking area based on user profile similarity. Int J Eng Technol 8(6):458–461

[14] Niu B, Li Q, Zhu X, Cao G, Li H (2014) Achieving k-anonymity in privacy-aware location-based services. In: INFOCOM, 2014 proceedings IEEE. IEEE, pp 754–762

[15] Alrahhal, M. S., Ashraf, M. U., Abesen, A., & Arif, S. (2017). AES-Route Server Model for Location based Services in Road Networks. International Journal Of Advanced Computer Science And Applications, 8(8), 361-368.

[16] Duckham M, Kulik L (2005) A formal model of obfuscation and negotiation for location privacy. In: International conference on pervasive computing. Springer, Berlin, pp 152–170

[17] Domingo-Ferrer J, Bras-Amoro´s M, Wu Q, Manjo´n J (2009) User-private information retrieval based on a peer-to-peer community. Data Knowl Eng 68(11):1237–1252

[18] Sen, A. A. A., Eassa, F. B., Yamin, M., & Jambi, K. (2018). Double Cache Approach with Wireless Technology for Preserving User Privacy. Wireless Communications and Mobile Computing, 2018.

[19] Shokri, R., Theodorakopoulos, G., Papadimitratos, P., Kazemi, E., & Hubaux, J. P. (2014). Hiding in the mobile crowd: Locationprivacy through collaboration. IEEE transactions on dependable and secure computing, 11(3), 266-279.

[20] Niu B, Li Q, Zhu X, Cao G, Li H (2015) Enhancing privacy through caching in location-based services. In: Computer communications (INFOCOM), 2015 conference. IEEE, pp 1017–1025

[21] Shin, K. G., Ju, X., Chen, Z., & Hu, X. (2012). Privacy protection for users of location-based services. IEEE Wireless Communications, 19(1).
[22] Yamin M, Sen AAA (2018) Improving privacy and security of user data in location based services. Int J Ambient Comput Intell (IJACI) 9(1):19–42

Abdullah M. Basahel
Associate Professor of MIS
Faculty of Economics and Administration, King Abdulaziz University, Jeddah, Saudi Arabia

Adnan Ahmed Abi Sen
Has MA in Web Sciences, MBA, and PhD in Computer Sciences, Faculty of Computer and Information Technology – KAU – Jeddah, Saudi Arabia

Professor Mohammad Yamin (ANU)
Professor of MIS
Faculty of Economics and Administration, KAU
King Abdulaziz University, Jeddah, Saudi Arabia

Sarah Alqahtani - Has MA degree from Marymount University.