Contact Tracing over Uncertain Indoor Positioning Data (Extended Abstract)

Tiantian Liu*, Huan Li†, Hua Lu‡, Muhammad Aamir Cheema§, Harry Kai-Ho Chan¶

*Department of Computer Science, Aalborg University, Denmark, liutt@cs.aau.dk

†College of Computer Science and Technology, Zhejiang University, China, lihuan.cs@zju.edu.cn

‡Department of People and Technology, Roskilde University, Denmark, luhua@ruc.dk

§Faculty of Information Technology, Monash University, Australia, aamir.cheema@monash.edu

¶Information School, University of Sheffield, United Kingdom, h.k.chan@sheffield.ac.uk

***Abstract*—Pandemics like COVID-19 often cause dramatic losses of human lives and societal impacts, urging efficient and effective contact tracing, especially in indoor venues where the risk of infection is higher. In this work, we formulate a novel query called Indoor Contact Query (ICQ) over raw, uncertain indoor positioning data that digitalizes people's indoor mobility. Given a query object o, e.g., a virus-carrying person, an ICQ analyzes uncertain indoor positioning data to find objects that most likely had close contact with o for a long period of time. To process ICQ, we propose a set of techniques. First, we design an enhanced indoor graph model to organize different types of data necessary for ICQ. Second, for indoor moving objects, we devise methods to determine uncertain regions and to derive positioning samples missing in the raw data. Third, we propose a query processing framework with a close contact determination method, a search algorithm, and multiple acceleration strategies. We conduct extensive experiments on synthetic and real datasets, which verify the efficiency and effectiveness of our proposals.**

I. Introduction

In the last 20 years, many pandemics broke out all over the world. For example, the COVID-19 pandemic caused more than 772 million infections, including over 6 million deaths [1]. To contain the spread of an epidemic, several measures have been practised, e.g., wearing masks, keeping social distance, and vaccination. It is also of high importance to conduct efficient and effective contact tracing, i.e., finding people who had close contact with an infected one.

In this work, we forumlate the Indoor Contact Query (ICQ) over indoor positioning data that digitalizes peoples' indoor mobility. Given a query object o, e.g., a person infected with coronavirus, an ICQ analyzes historical indoor positioning data to find those objects (other people) that most likely had close contact with o for a sufficiently long period of time. Processing ICQ faces significant technical difficulties. First, low-quality raw indoor positioning data makes it difficult to accurately determine uncertainty regions. Second, the complex indoor environment makes it difficult to model and process indoor data. Third, constraints in ICQ, such as distance and contact duration, render complex query processing.

Existing solutions fall short in solving ICQ. Proximity-based contact tracing [2], [3] do not consider spatial information and thus fail to customize close contact criteria. Location-based contact tracing methods [4], [5] focus on outdoor spaces and neglect indoor topology. Studies using indoor positioning data [6]–[9] ignore uncertainty of positioning data [6], fail to analyze the relationship between moving objects [7], [8], or focus on a different, online warning problem [9].

The main contribution of this paper is summarized as follows. (1) We formulate the Indoor Contact Query (ICQ) for contact tracing over historical uncertain indoor positioning data. (2) To adapt to the complex indoor environment and the low quality indoor positioning data, we build technical foundations of ICQ, including an enhanced indoor graph model, the determination of indoor uncertainty regions for moving objects, and the generation of derived samples from uncertain positioning data. (3) We propose an ICQ processing framework, together with a close contact determination method, a search algorithm, and multiple acceleration strategies. (4) We conduct extensive experiments which verify the efficiency and effectiveness of our proposals.

II. Problem Formulation

In a typical setting, an indoor positioning system aperiodically reports an indoor moving object o's positioning record as $\psi = (l, t, et)$, indicating o was found at location l at time t and the result expiring at a later time et. A location l is captured as (x, y, f), meaning a point (x, y) on a floor f. Each object o's **raw trajectory** Ψ_o is a time-ordered sequence of its positioning records. In Ψ_o, two consecutive positioning records $\psi_i = (l_i, t_i, et_i)$ and $\psi_{i+1} = (l_{i+1}, t_{i+1}, et_{i+1})$ satisfy that $et_i < t_{i+1}$. Thus, Ψ_o most often do not fully disclose the object whereabouts during Ψ_o's lifespan. To enable the comparison of two objects' trajectories throughout a common part of their lifespans, we generate the **sampled trajectory** Ψ_o^s of each object, where all objects' location samples are aligned in the time dimension. The details can be found in [10]. Ψ_o^s is a time-ordered sequence of the sample sets at all sampling times, formally $\Psi_o^s = \langle (S_1, t_1^s), \ldots, (S_W, t_W^s) \rangle$, where S_w is a set of samples $s_i = (l_i, \rho_i)$ and t_w^s is the sampling time. Each sampled location l_i is associated with a probability ρ_i.

Definition 1 (Contact Distance)**.** *The contact distance between two indoor locations l_i and l_j is computed as*

$$dist(l_i, l_j) = \begin{cases} ||l_i, l_j||_E, & \textit{if } l_i \textit{ and } l_j \textit{ fall in the same partition;} \\ \infty, & \textit{otherwise.} \end{cases} \quad (1)$$

where $||\cdot, \cdot||_E$ computes the Euclidean distance.

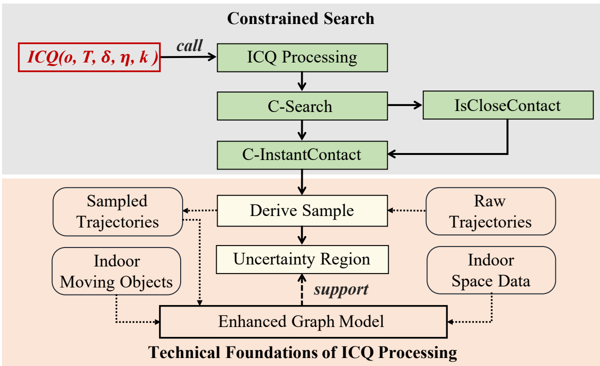


Fig. 1: The solution framework for indoor contact query

Here, a *partition* refers to a basic topological unit of the indoor space, e.g., rooms or staircases. Two objects are not considered contact if they are in different partitions[1].

Definition 2 (Contact Probability). *Given a distance threshold δ and two indoor objects o_i and o_j, we obtain their sample sets at time t^s_w as $S_{i,w}$ and $S_{j,w}$. The contact probability*

$$\mathsf{P}(o_i, o_j, t^s_w) = \sum_{(s_i,\rho_i)\in S_{i,w},(s_j,\rho_j)\in S_{j,w}} c(l_i, l_j)\rho_i\rho_j, \quad (2)$$

where $c(l_i, l_j)$ is 1 if the $dist(l_i, l_j) \leq \delta$ and 0 otherwise.

We say two objects are **instant contact** with each other at time t if their contact probability is higher than a specified threshold η. Our research problem is formulated as follows.

Problem (Indoor Contact Query, ICQ). *Given a query object o, a time interval T, a distance constraint δ, a contact probability threshold η, and a contact number k, an indoor contact query* ICQ(o, T, δ, η, k) *returns all moving objects having instant contacts with o for at least k consecutive sampling times within T.*

III. Solution Framework

Fig. 1 shows our solution for ICQ. The bottom layer lays the technical foundations of ICQ processing, where an enhanced graph model is constructed to accommodate the indoor space data, indoor moving objects, and sampled trajectories for relevant computations. On top of the enhanced graph model, ICQ first finds a moving object's uncertainty region, i.e., the object's possible locations at a sampling time. Based on that, ICQ derives samples that are to be maintained in the enhanced graph model for subsequent use.

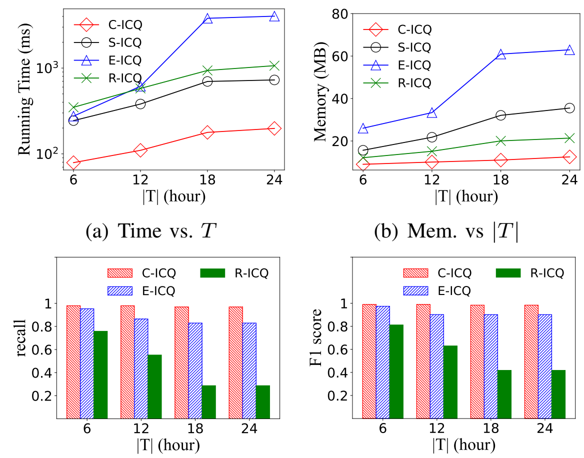
To process a query instance ICQ(o, T, δ, η, k), an overall framework is proposed as *ICQ Processing*. It finds all objects in close contact with the query object o by calling the constrained search *C-Search*, which employs a set of strategies for acceleration. Then, *C-Search* calls *C-InstantContact* to determine the instant close contact and further help set the start timestamp. *C-Search* also calls *IsCloseContact* for close contact determination when processing a candidate object. The *IsCloseContact* calls *C-InstanceContact* to compute the contact probability at an instant timestamp.

[1]Our method can be easily adapted to cover the case that consider objects located near a door but in separate partitions as in contact.

IV. Experiments

Baselines. As no existing methods can solve the indoor contact query, we design three baselines. S-ICQ sequentially processes each sampling time and computes the concrete contact probabilities. E-ICQ finds the uncertainty region at each unseen sampling timestamp based on Euclidean distance. R-ICQ searches for close contact objects over raw trajectories. **Datasets**. We use both real and synthetic datasets. For synthetic datasets, we generate a 5-floor indoor space that is of 1,368 × 1,368 m^2 with 141 partitions and 216 doors on each floor. We simulate raw trajectories of moving objects in the indoor space for 24 hours. More details can be found in [10].

Fig. 2 shows the results of varying $|T|$. We omit S-ICQ in the effectiveness studies as S-ICQ and C-ICQ return the same results. For all four measures, C-ICQ performs best.



(a) Time vs. T (b) Mem. vs $|T|$

(c) Recall vs $|T|$ (d) F1-score vs $|T|$

Fig. 2: Effect of $|T|$

Acknowledgement

This work was supported by Independent Research Fund Denmark (No. 8022-00366B) and Australian Research Council (DP230100081 and FT180100140).

References

[1] https://covid19.who.int, accessed: 2023-11-30.
[2] R. L. Rivest, D. Weitzner, L. Ivers, I. Soibelman, and M. Zissman, "Pact: Private automated contact tracing," *https://pact.mit.edu/*, 2020.
[3] P. C. Ng, P. Spachos, S. Gregori, and K. N. Plataniotis, "Epidemic exposure tracking with wearables: A machine learning approach to contact tracing," *IEEE Access*, vol. 10, pp. 14 134–14 148, 2022.
[4] X. Zhang, S. Ray, F. Shoeleh, and R. Lu, "Efficient contact similarity query over uncertain trajectories," in *EDBT*, 2021, pp. 403–408.
[5] K. Li, L. Chen, S. Shang, Y. Liu, H. Wang, P. Kalnis, and B. Yao, "Towards controlling the transmission of diseases: Continuous exposure discovery over massive-scale moving objects," in *IJCAI*, 2022, pp. 3891–3897.
[6] H. Lu, B. Yang, and C. S. Jensen, "Spatio-temporal joins on symbolic indoor tracking data," in *ICDE*. IEEE, 2011, pp. 816–827.
[7] H. Li, H. Lu, L. Shou, G. Chen, and K. Chen, "In search of indoor dense regions: An approach using indoor positioning data," *IEEE TKDE*, vol. 30, no. 8, pp. 1481–1495, 2018.
[8] ——, "Finding most popular indoor semantic locations using uncertain mobility data," *IEEE TKDE*, vol. 31, no. 11, pp. 2108–2123, 2018.
[9] H. K.-H. Chan, H. Li, X. Li, and H. Lu, "Continuous social distance monitoring in indoor space," *Proc. VLDB Endow.*, vol. 15, no. 7, pp. 1390–1402, 2022.
[10] T. Liu, H. Li, H. Lu, M. A. Cheema, and H. K.-H. Chan, "Contact tracing over uncertain indoor positioning data," *IEEE TKDE*, vol. 35, no. 10, pp. 10 324–10 338, 2023.