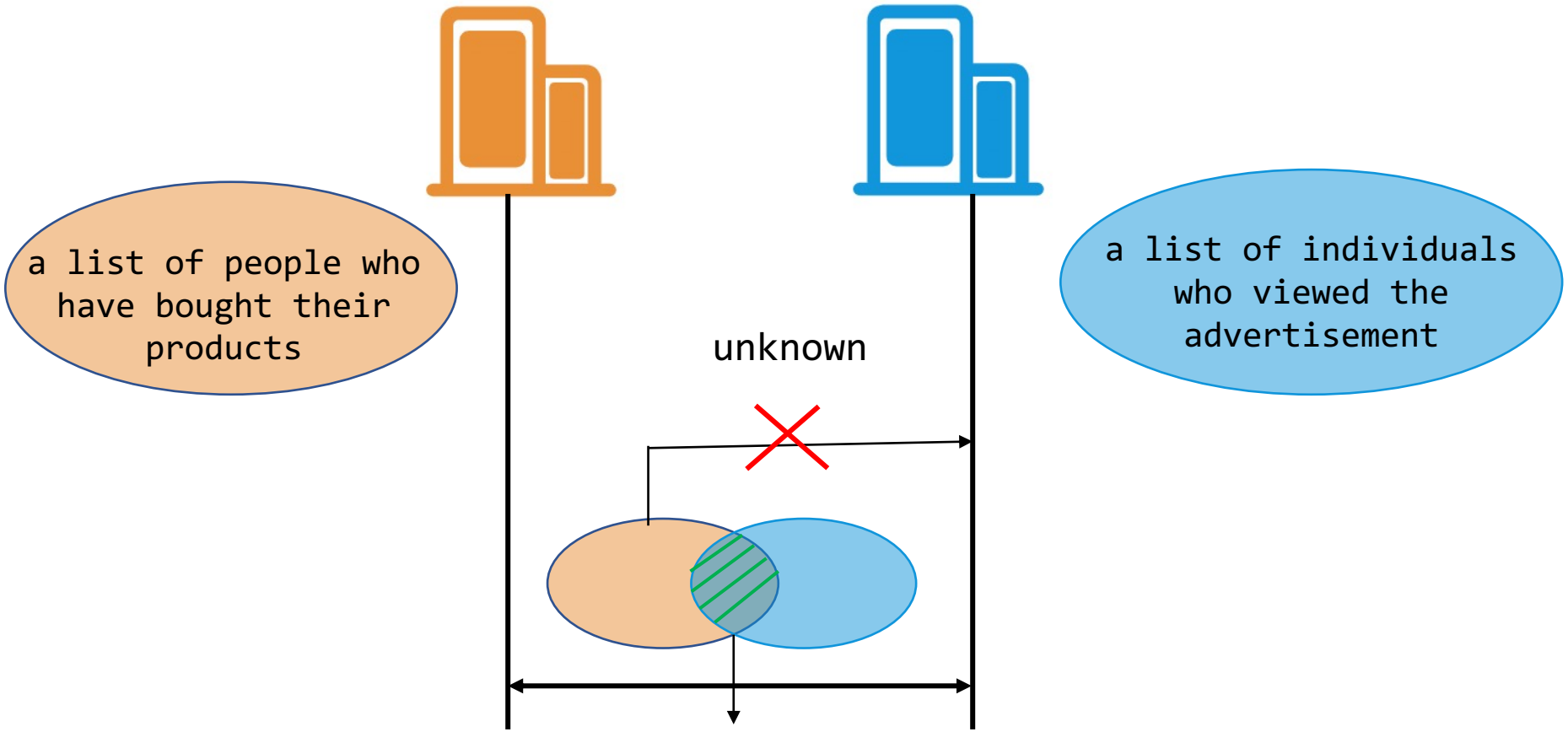
PSI based on ECDH

Presenter: Wenting Chang from Alipay

Use case 1: Joint marketing

Company A and Company B have their own customer lists and want to work together on a joint marketing campaign to target customers who may be interested in both companies. Using PSI technology, two companies can identify the intersecting customer base.



Using ECDH-PSI, the effect of an advertisement can be accurately calculated, without revealing any extra information about the users.

Use case 2: Promotion of bank card user activation

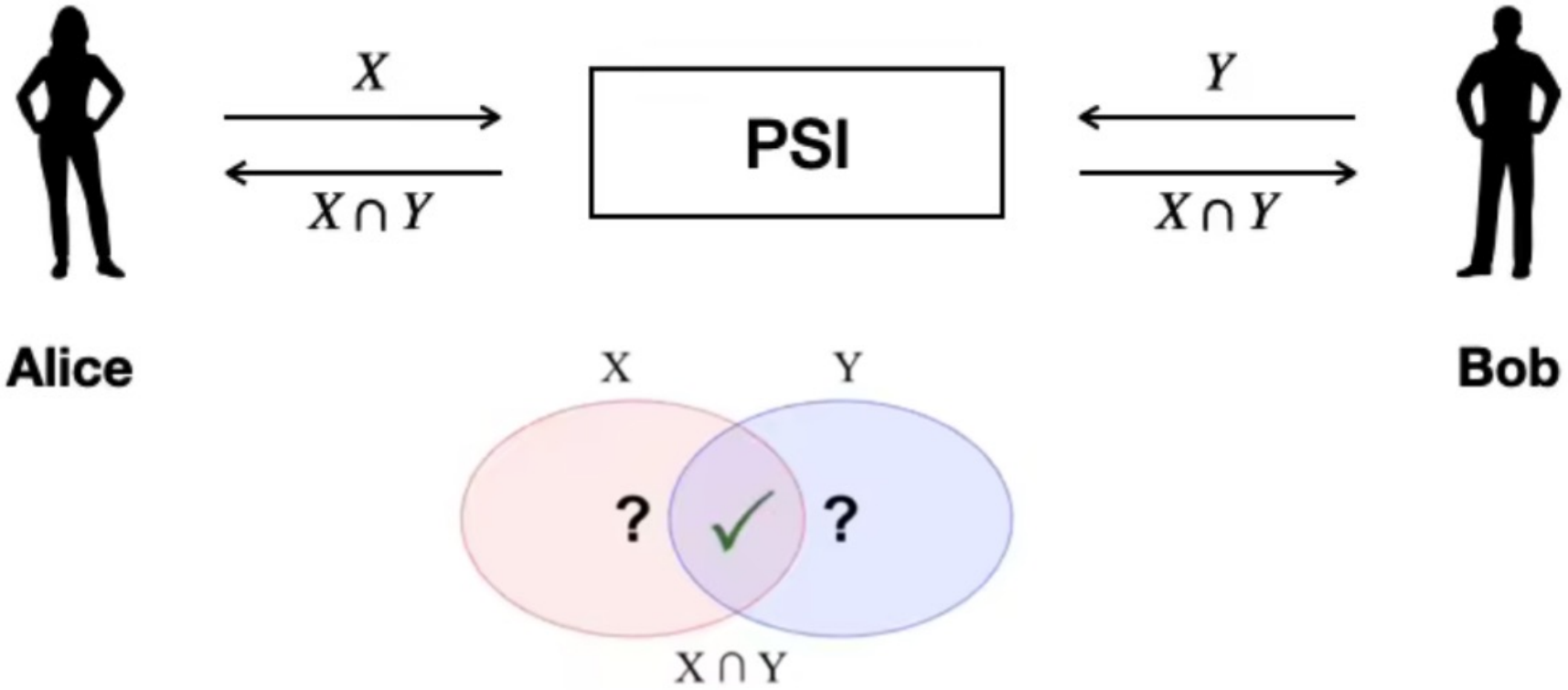
Banks implement various initiatives to attract new users to open bank cards and make their first transactions. For inactive bank card users, banks will focus on popular payment scenarios such as grocery and dining expenditures, offering payment discounts to encourage users to transition to bank cards for transactions.

Reach out online shopping customers to use bank card

offer discounts to encourage users to pay money with bank card

Banks perform PSI with e-commerce and other platforms to identify potential targets for user activation.

Private Set Intersection (PSI) enables between two or more parties to **identify overlapping data elements while preserving the confidentiality of each party's complete dataset.**



- **A protocol** to negotiate required parameters and exchange data
- The negotiation and data **message structure** to realize grammar and semantic consistency

PSI algorithm comparison

PSI algorithm	Characteristic
HASH-based PSI	• higher performance • lower security, e.g. unilateral exhaustive attack
Elliptic Curve Diffie-Hellman (ECDH) based PSI	• limited communication bandwidth • moderate computation requirement • be capable for handling data at the scale of billions
Garbled Circuit based PSI	• higher communication bandwidth • higher computation requirement • Higher design complexity
oblivious transfer (OT) based PSI [1)]	• higher communication bandwidth • lower computation requirement
Vole based PSI [2)]	• algorithm is continuously iterative • lower communication bandwidth • easily construct malicious security PSI
Fully Homomorphic Encryption based PSI [3)]	• higher computation requirement • lower communication bandwidth • fully homomorphic algorithm is still evolving

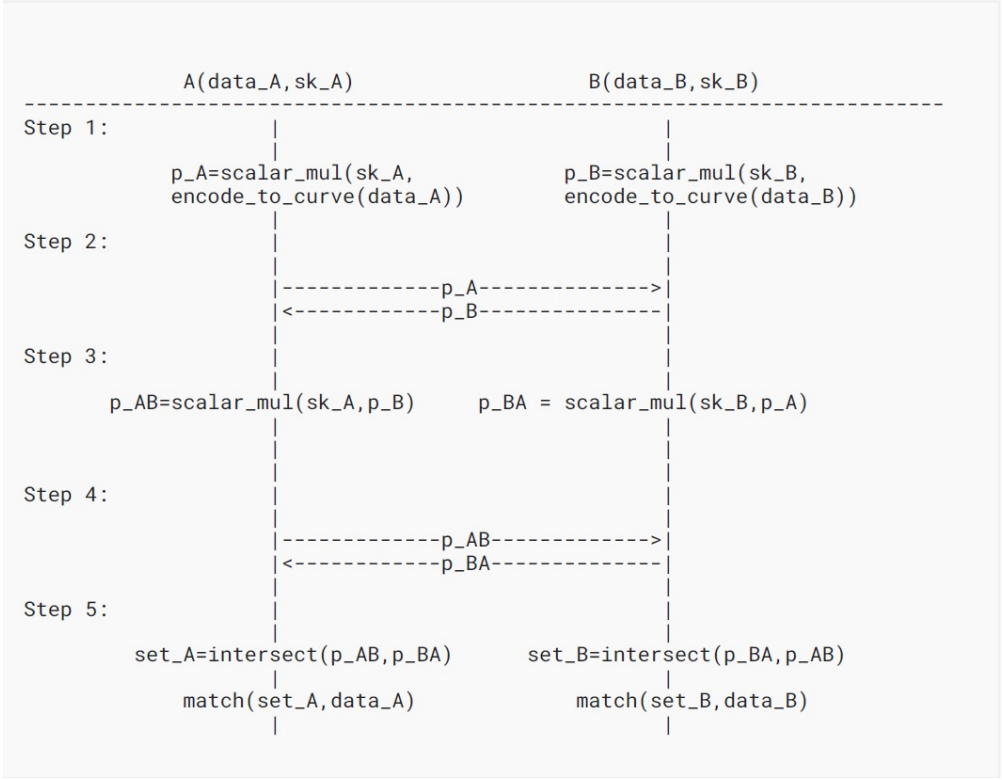
This contribution focus on ECDH-PSI algorithm, since it's readily deployable with limited bandwidth demands and strong security, and the performance can be improved by multithreading and distributed optimization

1) Vladimir Kolesnikov, Ranjit Kumaresan, Mike Rosulek, Ni Trieu. Efficient Batched Oblivious PRF with Applications to Private Set Intersection
2) S.Raghuraman, P. Rindal. Blazing Fast PSI from Improved OKVS and Subfield VOLE
3) Hao Chen, Zhicong Huang, Kim Laine, Peter Rindal. Labeled PSI from Fully Homomorphic Encryption with Malicious Security,

Overview of ECDH-PSI algorithm

In ECDH-PSI, both participants agree on a Elliptic Curve group parameter `G` and generate ECDH key pairs over `G`. The keys are then used to mask the original data with scalar multiplications.

- Step1: A participant maps its data items to points over an elliptic curve with `encode_to_curve`, and mask the points locally with its own private key by `scalar_mul`.
- Step2: A and B exchange their locally masked data as EC points.
- Step3: Upon receiving the masked data from its partner, a participant doubly masked the received points with its private key.
- Step4: A participant sends the doubly-masked points back to its partner.
- Step5: The participant calculates intersection of the set calculated in Step 3 and the set received in Step 4, and finally outputs the original data corresponding to the intersection.



Overview of ECDH-PSI

- **Proposal:** https://datatracker.ietf.org/doc/draft-ecdh-psi/
- **Code:** https://github.com/secretflow/interconnection

Thanks!

Q&A

Contact: wenting.chang@antgroup.com

tianwu.wyc@antgroup.com