2020 embedded VISION summit®

Multi-modal Re-Identification: IOT + Vision for Residential Community Tracking

Dr. Kit Thambiratnam
General Manager, AI Center, The Seedland Group

Sept. 2020

SEEDLAND

Outline

- **Landscape** - Residential Communities in China
- **Problem** – Why tracking in communities?
- **Multi-modal tracking** - framework and results
  - Tracklets
  - Identity Stamping
  - Identity Conflation
- **Conclusions**

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Residential Communities in China

Residential communities focus on the **needs of households**, providing secure housing integrated with recreation, groceries, retail and lifestyle services.

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

The Seedland Group

Craftmanship creates quality life products

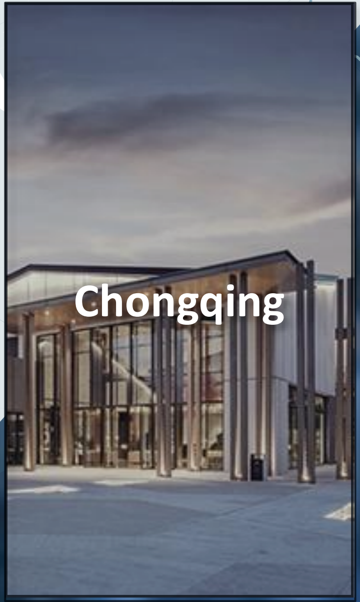
For 14 years, the Seedland Real Estate Group Co., Ltd. has been committed to the exploration and innovation of human science and technology in all aspects of life, connecting science and technology with humanities, and redefining human understanding of the relationship between themselves and living space.

31 cities

45 projects

6 fastest city clusters

Guangzhou


Chongqing

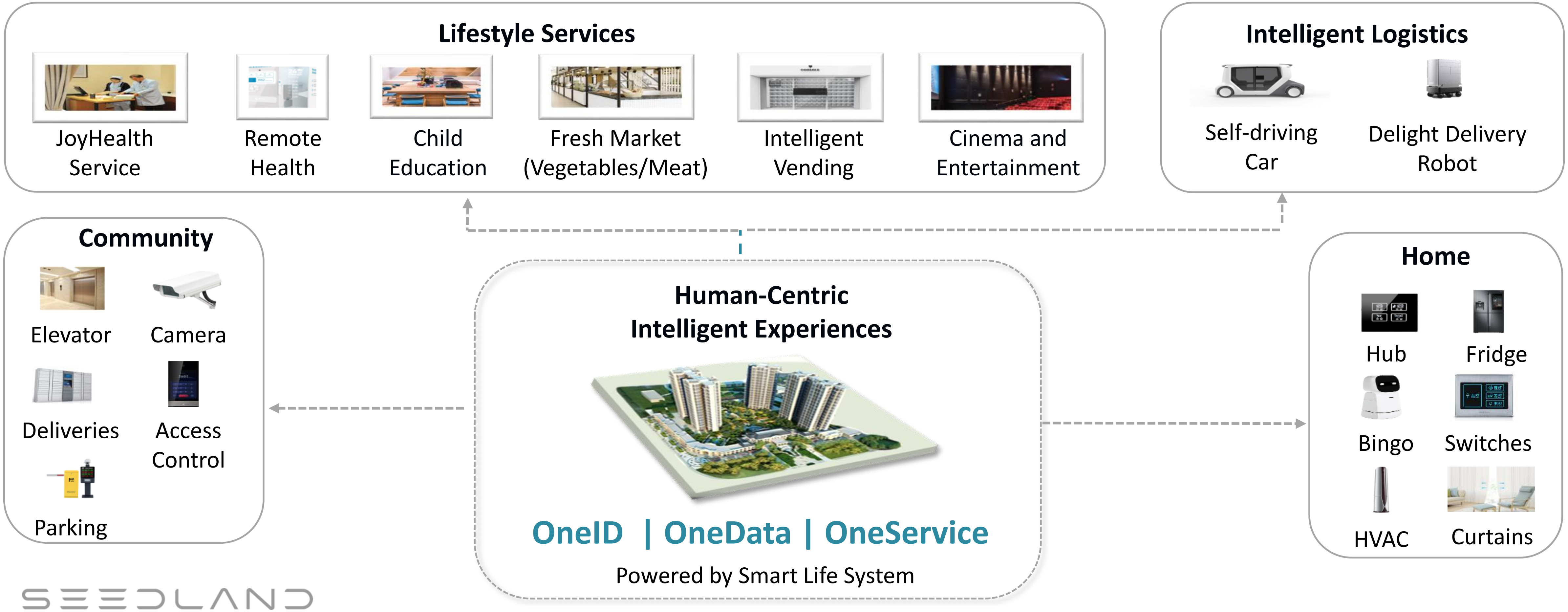
Qingdao

Shenzhen

n may be re

Seedland Smart Community Ecosystem

Our goal is to pair smart **home**, smart **community**, signature smart premium **lifestyle services** and intelligent **logistics** to provide a **premier residential lifestyle experience**.



Lifestyle Services

JoyHealth Service

Remote Health

Child Education

Fresh Market (Vegetables/Meat)

Intelligent Vending

Cinema and Entertainment

Intelligent Logistics

Self-driving Car

Delight Delivery Robot

Community

Elevator

Camera

Deliveries

Access Control

Parking

Human-Centric Intelligent Experiences

OneID | OneData | OneService

Powered by Smart Life System

Home

Hub

Fridge

Bingo

Switches

HVAC

Curtains

SEEDLAND

No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Smart Life System

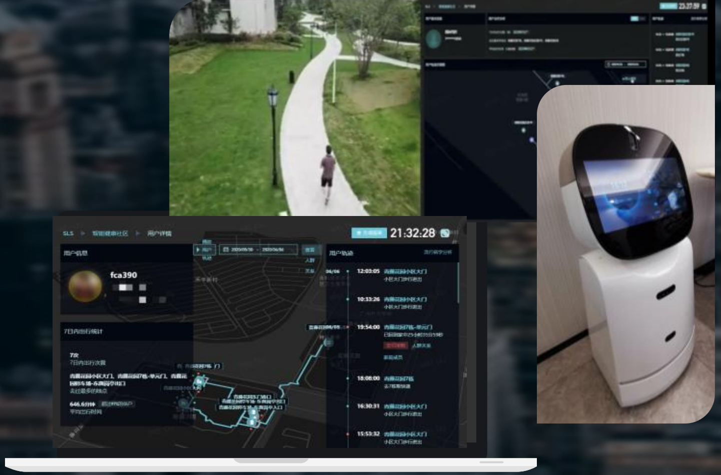
Community, Health, Business

THREE SLS SOLUTIONS

Operating in Guangzhou Ivy

SLS Smart Community Solutions

Intelligent security, child safety, people flow analysis, intelligent delivery and contactless access control based on real-time IOT, cameras and AI.

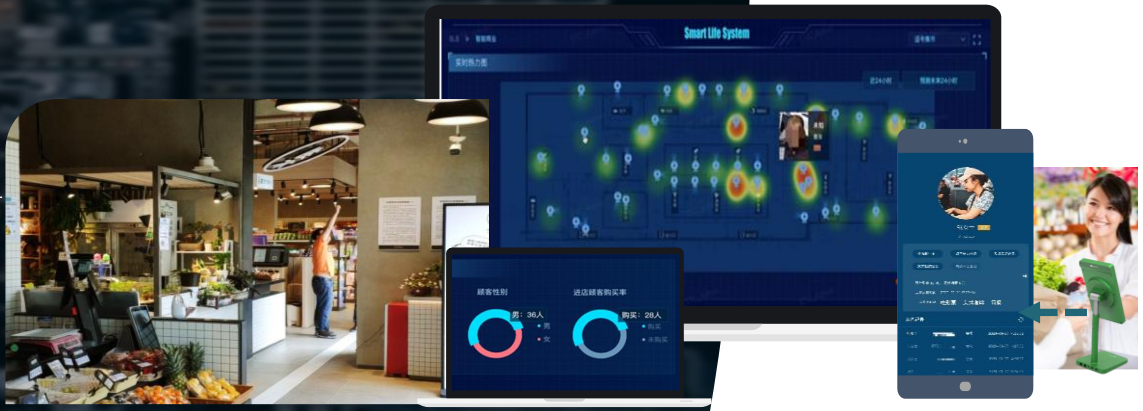


SLS Intelligent Community Health

Intelligent health situational back-tracing for emergency contact tracing and multi-level risk alerts.

SLS Smart Business Services

Personalized and intelligent shopping through integrated community-wide ordering, standardized delivery, AI customer service assistance, intelligent business operations analysis.



Community Tracking

Why Community Tracking?

Community tracking **delivers value** to **residents** and **property management** while strictly **preserving privacy** and allowing **user opt-out**.

Playground Tracing

Only residents/approved visitors enter child areas

Delivery Tracing

Delivery workers do not visit unnecessary areas.

Prohibited Tracing

Tracing rule breakers for property management

Contact Tracing

Contact tracing for non-phone users (kids/elderly)

Suspicious Tracing

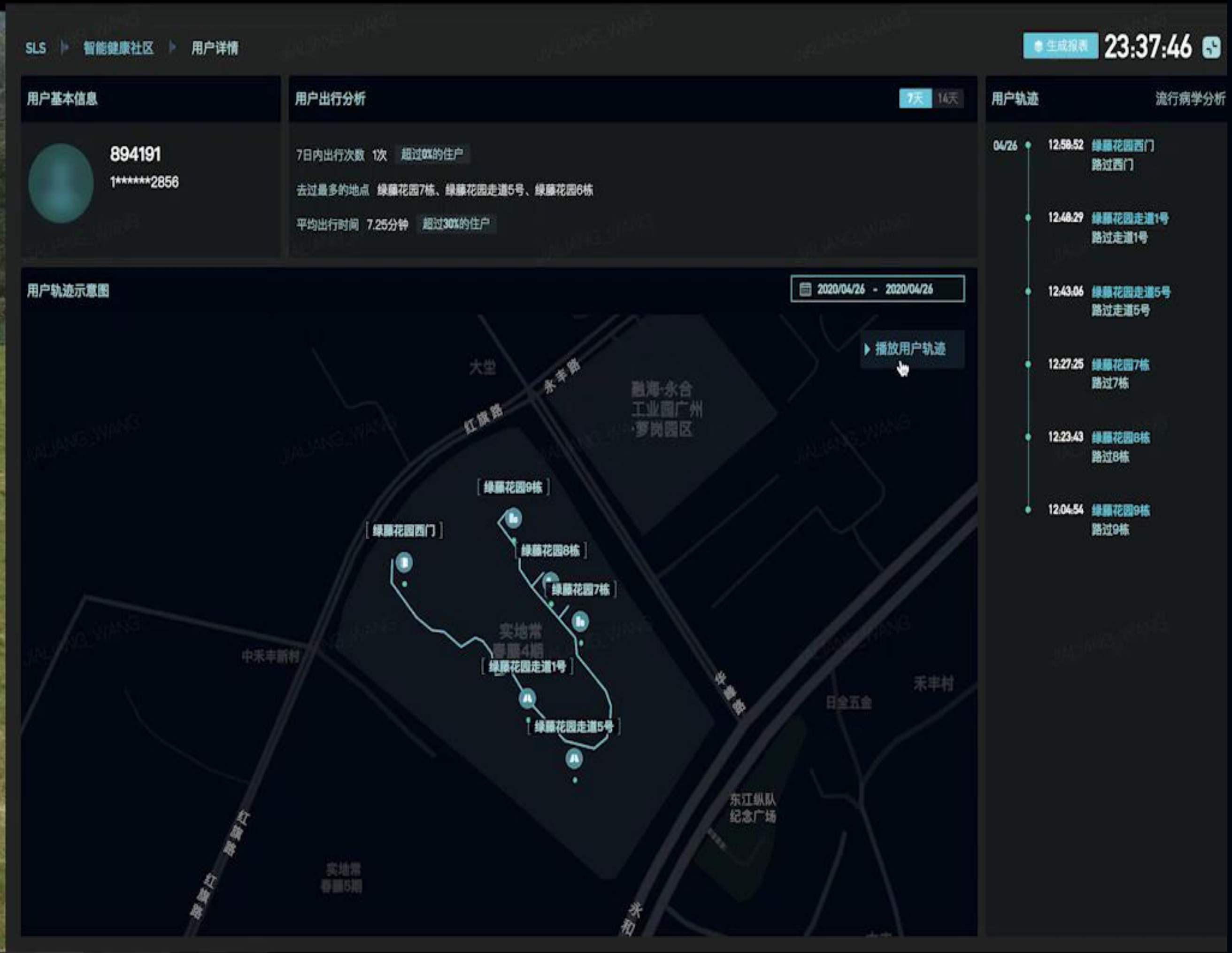
Retracing paths of suspicious behavior

Lost Item Tracing

Retracing to find lost items, phones, toys, …

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Emergency Contact Back-Tracing



SLS ▸ 智能健康社区 ▸ 用户详情

生成报表 23:37:46

用户基本信息

894191
1******2856

用户出行分析 7天 14天

7日内出行次数 1次 超过0%的住户

去过最多的地点 绿藤花园7栋、绿藤花园走道5号、绿藤花园6栋

平均出行时间 7.25分钟 超过30%的住户

用户轨迹示意图

2020/04/26 - 2020/04/26

▸ 播放用户轨迹

用户轨迹 流行病学分析

04/26
12:58:52 绿藤花园西门
路过西门

12:48:29 绿藤花园走道1号
路过走道1号

12:43:06 绿藤花园走道5号
路过走道5号

12:27:25 绿藤花园7栋
路过7栋

12:23:43 绿藤花园8栋
路过8栋

12:04:54 绿藤花园9栋
路过9栋

Combining Multiple Modalities

A **user** will **interact** with **multiple devices** and **multiple modalities** in their journey.
Multi-modal allows us **to see beyond cameras** and **generate semantic knowledge**.

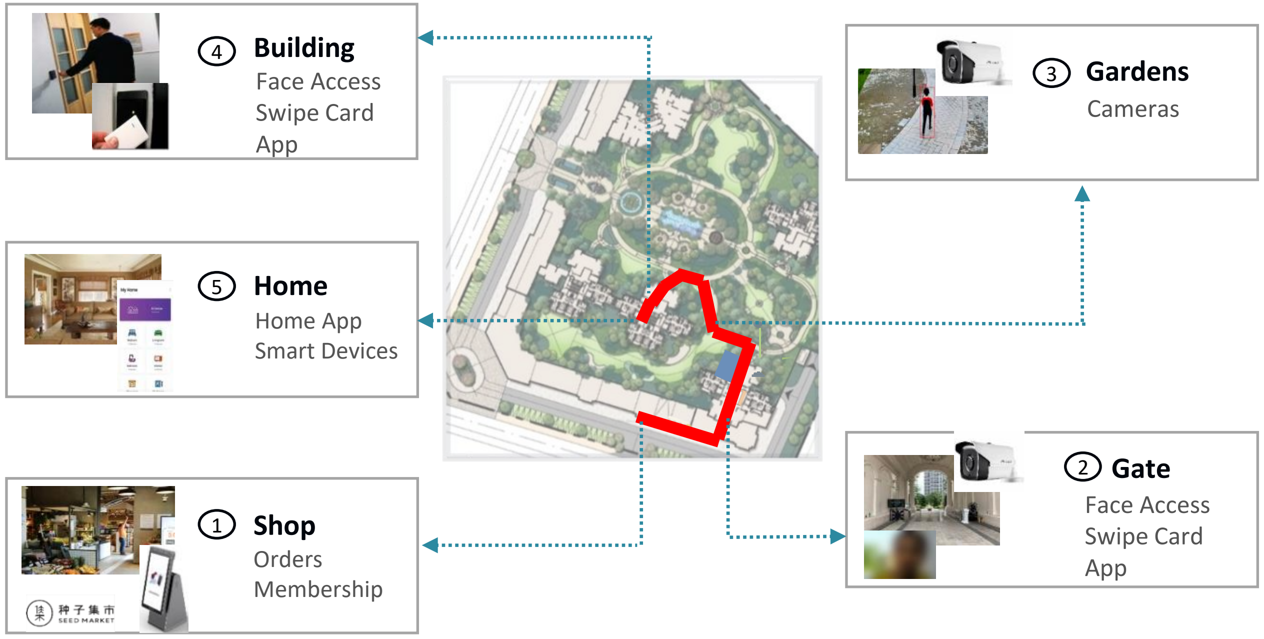
1.Family Structure

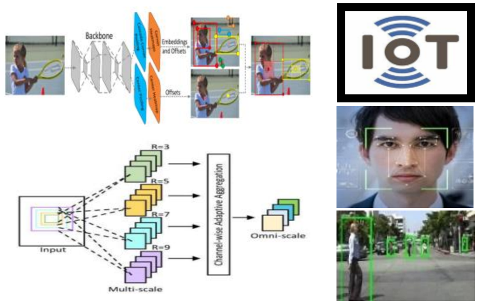
2. Travel Habits

3. Consumption

4. Health

5. Community Graph





Multi-Modal Tracking

Detection (CenterNet '19)
Vector Search (ANN)
ReID (OSNet 'CVPR19)
Face Attrib. (ArcFace CVPR19)
IOT Conflation (Search)

Family -- Travel -- Consumption

Semantic Knowledge Graph
Docs
Video
Content
Metadata
Proprietary Master Data
Commercial Master Data
LOD

Health -- Community Graph

Knowledge Graph

Real-time knowledge fusion
Real-time tagging (GBDT)
Real-time graph (RT Graph)
Relationship mining (Entropy)

No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Challenges of Tracking In Residential Communities

Face-Based Identity

- Bias - Face recognition is known to perform poorly on children and elderly.
- Angle - High-positioned security cameras make recognition difficult.

Device-Based Identity

- Mobile phones are NOT everywhere - children and elderly.
- Bluetooth-tracking low penetration in China.

Privacy

- Must support do-not-track.
- Must support usage for intended purposes only.

© 2020 Seedland

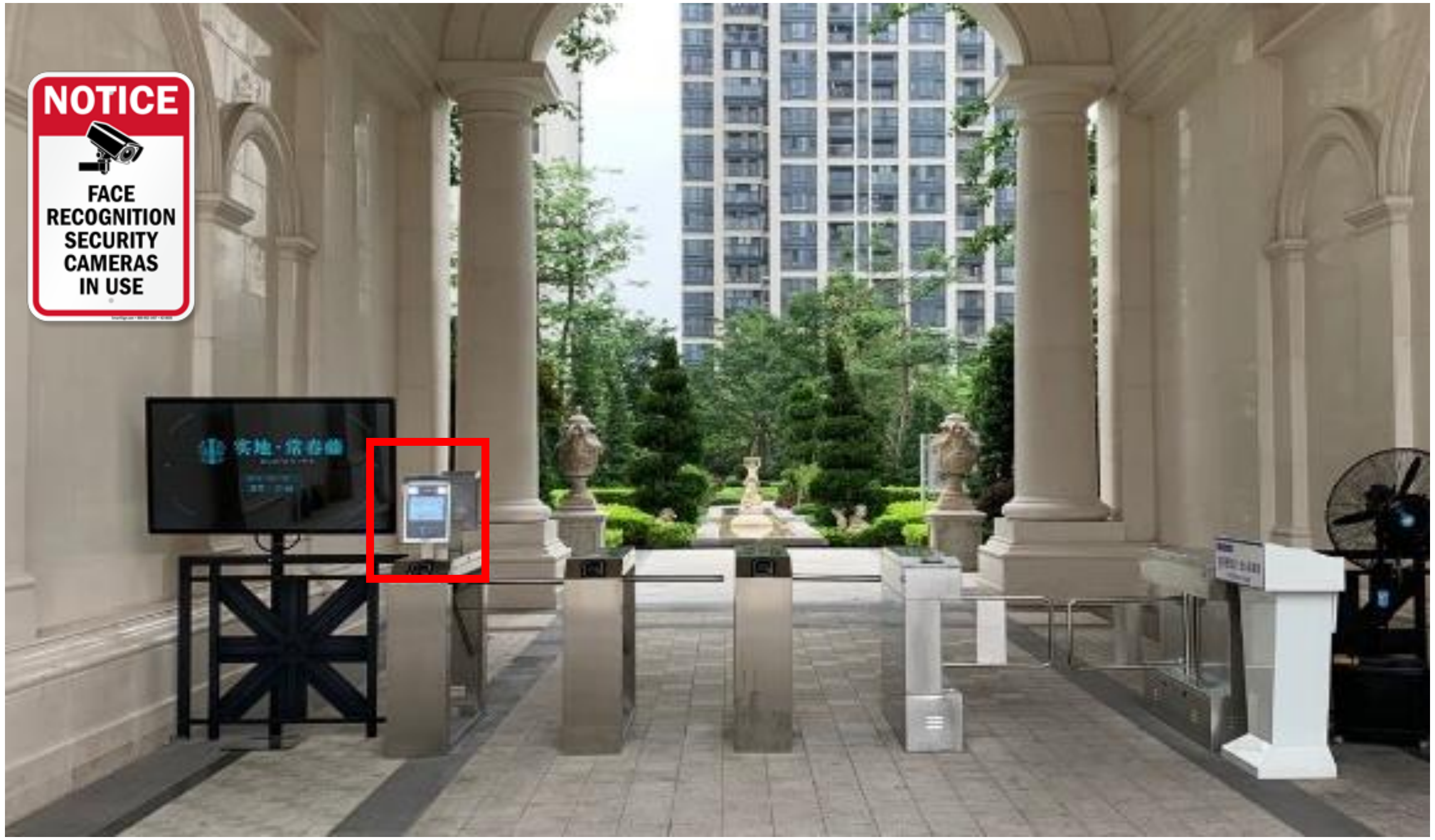
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

User Consent and Tracking

Obtain **explicit consent** during smart gate registration

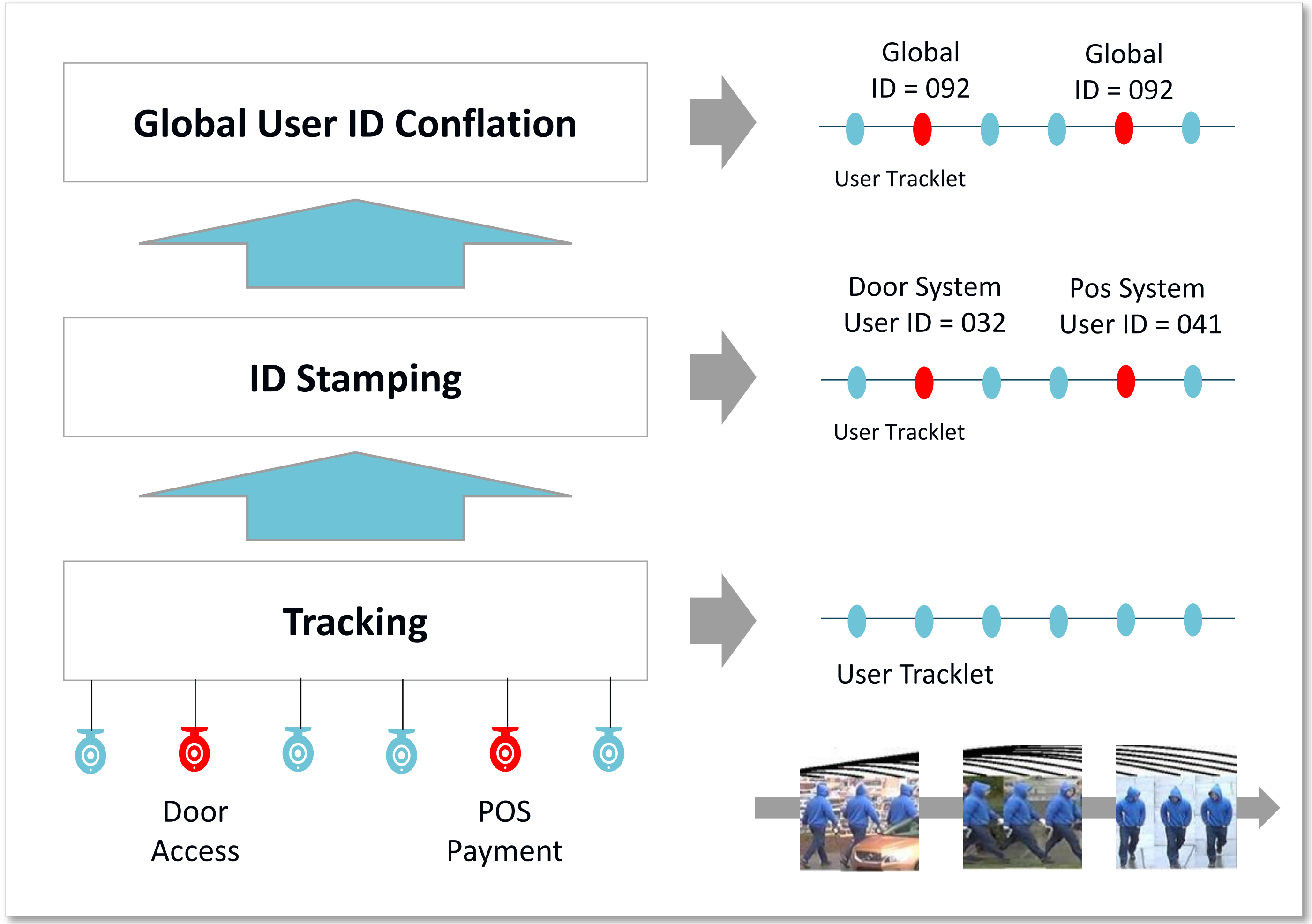
Consent agreement clearly lists **specific user-value features** – no blanket permission for tracking

Non-ID associated short-term tracking (eg. suspicious person back-tracing) are **implicitly agreed** to via **security notice**



© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Cross-Day Stable User ID Tracking



Key Challenges

- ID stable across days
- Real-time
- Cost-efficient
- Visual invariance (clothes, bags, hats, …)
- Population bias (children, elderly).

Leverage multiple modalities while real-time and cost efficient.

© 2020 Seedland

No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

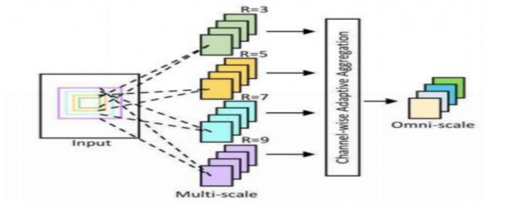
Tracking

SEEDLAND

Cross-Camera Tracking (REID)

Key idea: Recognize same person across cameras by searching a continuously updated database of previously-seen entities

OSNet[1] **multi-scale feature** learning robust to scale, camera-distance

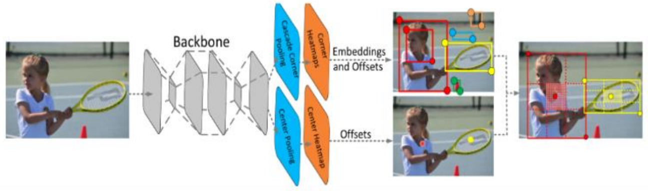


Tracklet ID - Assign to entire track
(smooth per-frame noisy recognition)

t	2	3	4	5	6	7
id	014	014	014	012	014	012
m	0.8	0.9	0.9	0.7	0.5	0.8

Camera Group → **ID-Aware Tracking** → **Omni-Scale Embedding** → **Cross-Camera ReID Search** → **Tracklet ID Assign** → **Seen Entity Update** → **Seen Entities** → **Cross-Camera ReID Search**

Camera Group



ID-aware tracking dramatically improves tracking MOTA and high FPS

CenterNet (detection) +
Kalman (trajectory) +
Embedding Distance (identity)

Seen Entities

id	camera	exemplars	m
013	c41	[...]	0.8
013	c44	[...]	0.9

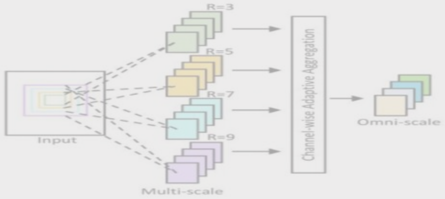
Detect **new entities**
Add **new exemplars** for known entities

[1] Zhou, Kaiyang and Yang, Yongxin and Cavallaro, Andrea and Xiang, Tao, “Omni-Scale Feature Learning for Person Re-Identification”, in The IEEE International Conference on Computer Vision (ICCV), 2019

Cross-Camera Tracking (REID)

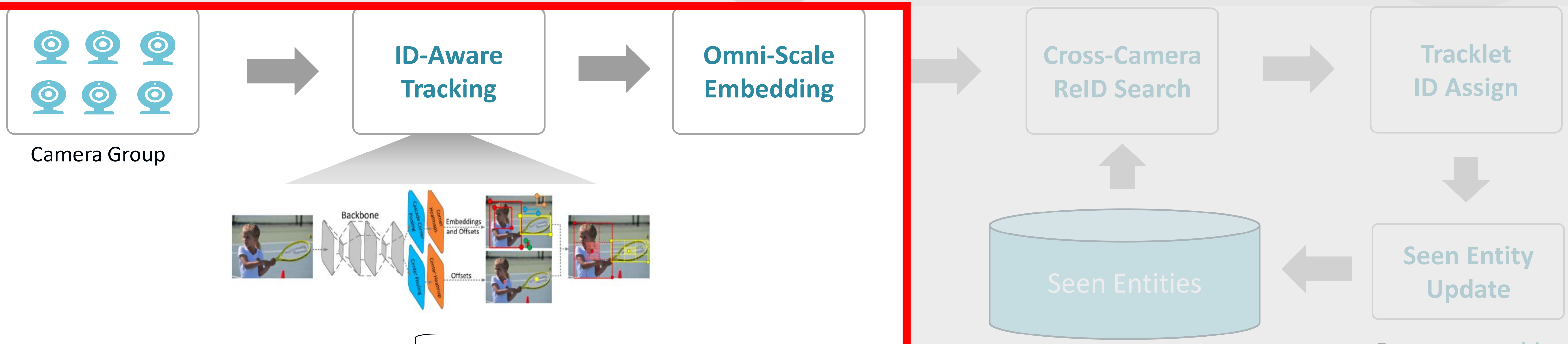
Key idea: Recognize same person across cameras by searching a continuously updated database of previously-seen entities

OSNet[1] **multi-scale feature** learning robust to scale, camera-distance



Tracklet ID - Assign to entire track
(smooth per-frame noisy recognition)

t	2	3	4	5	6	7
id	014	014	014	012	014	012
m	0.8	0.9	0.9	0.7	0.5	0.8



ID-aware tracking dramatically improves tracking MOTA and high FPS

CenterNet (detection) +
Kalman (trajectory) +
Embedding Distance (identity)

id	camera	exemplars	m
013	c41	[...]	0.8
013	c44	[...]	0.9

Detect **new entities**
Add **new exemplars** for known entities

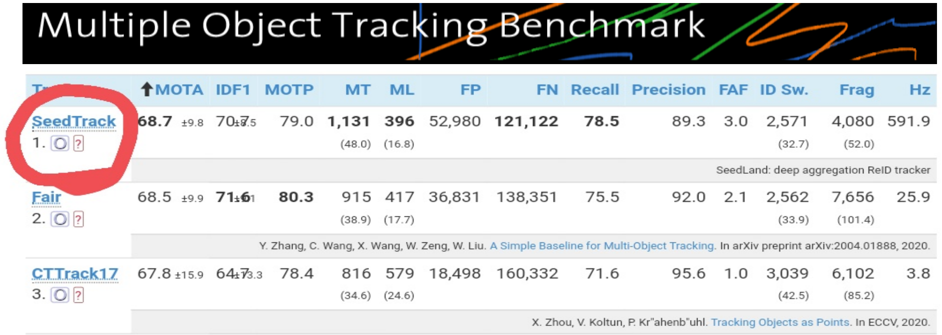
[1] Zhou, Kaiyang and Yang, Yongxin and Cavallaro, Andrea and Xiang, Tao, “Omni-Scale Feature Learning for Person Re-Identification”, in The IEEE International Conference on Computer Vision (ICCV), 2019

Tracking Tasks and Public Benchmarks

Single-Camera Tracking – Detect and track a person within one camera

MOT16

MOT20



Multiple Object Tracking Benchmark

Tr	↑MOTA	IDF1	MOTP	MT	ML	FP	FN	Recall	Precision	FAF	ID Sw.	Frag	Hz
SeedTrack 1. ○ ?	**68.7** ±9.8	70.7 ±8.5	79.0	**1,131** (48.0)	**396** (16.8)	52,980	**121,122**	**78.5**	89.3	3.0	2,571 (32.7)	4,080 (52.0)	591.9
	SeedLand: deep aggregation ReID tracker												
Fair 2. ○ ?	68.5 ±9.9	**71.6** ±6.1	**80.3**	915 (38.9)	417 (17.7)	36,831	138,351	75.5	92.0	2.1	2,562 (33.9)	7,656 (101.4)	25.9
	Y. Zhang, C. Wang, X. Wang, W. Zeng, W. Liu. A Simple Baseline for Multi-Object Tracking. In arXiv preprint arXiv:2004.01888, 2020.												
CTTrack17 3. ○ ?	67.8 ±15.9	64.7 ±13.3	78.4	816 (34.6)	579 (24.6)	18,498	160,332	71.6	95.6	1.0	3,039 (42.5)	6,102 (85.2)	3.8
	X. Zhou, V. Koltun, P. Kr"ahenb"uhl. Tracking Objects as Points. In ECCV, 2020.												

Cross-Camera Tracking (REID) – Identify the same user across different cameras

Duke MTMC Set

MSMT17 Set

Method	Benchmark Set Top-1 Accuracy			Model Memory Size
	Market1501	DukeMTMC	MSMT17	
Seedland	95.1	88.6	78.7	11M
PCB	93.8	83.3	68.2	120M
BFE-Net	94.0	88.9	-	130M
DG-Net	94.8	83.6	77.2	108M

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Tracking is critical for down-stream understanding Learnings from ACM2020 Grand Challenge

- #1 position in the ACM 2020 Multimedia Grand Challenge for Large-scale Human-centric Video Analysis international competition.

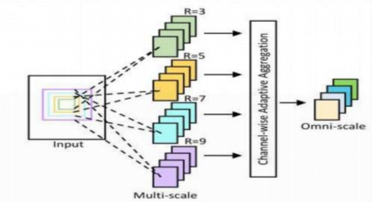
- High ranked in down-stream intelligence (pose tracking, action recognition) **primarily because of improved dense crowd tracking**

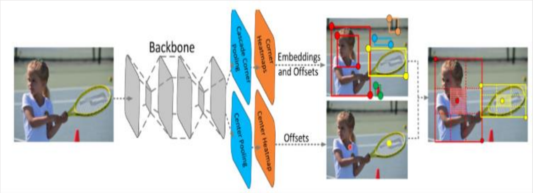
ID-aware tracking

Omni-scale embed

Multi-Stage model ensemble

Perspective/occlude augment





Team Name	# Institution	MOTA
[1st] Adaptive FairMOT	iSEE-SYSU & ACCUVISION	60.2282
[2nd] JiaRen.AI	Seedland	56.0474
[3rd] Crowd-Tracker	Xidian University	55.5163
Try private	Amazon	47.8120
NewTracker	Tencent	46.4815

Team Name	# Institution	w_AP@avg
[1st] Seedland.Tech	Seedland	57.5091
[2nd] ccc	YiTu & National University of Singapore	56.3375
[3rd] DH_IBA	Dahua Technology	55.1719
JDAI	JD Tech & UESTC	55.0139
WanDan	UESTC	54.5282

Team Name	# Institution	MOTA
[1st] Seedland.Tech	Seedland	63.9686
[2nd] Try	YiTu & National University of Singapore	61.7941
[3rd] SimpleTrack	Chinese University of Hong Kong	56.9834
DeepBlueAI	DeepBlueAI	55.1543
Commander_test4	XFORWARDAI	53.7671

Team Name	# Institution	mAP@avg
[1st] MSF	YiTu & National University of Singapore	0.2605
[2nd] VM	Seedland	0.2548
[3rd] CF	City University of Hong Kong	0.1531
8A	Tencent	0.1509
only_person_rgb	Sun Yat-sen University	0.1086

SEEDLAND

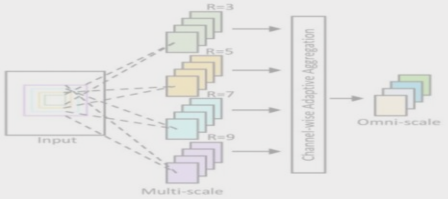
© 2020 Seedland

No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Cross-Camera Tracking (ReID)

Key idea: Recognize same person across cameras by searching a continuously updated database of previously-seen entities

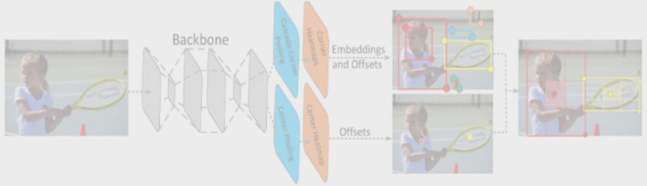
OSNet[1] **multi-scale feature** learning robust to scale, camera-distance



Tracklet ID - Assign to entire track
(smooth per-frame noisy recognition)

t	2	3	4	5	6	7
id	014	014	014	012	014	012
m	0.8	0.9	0.9	0.7	0.5	0.8

Camera Group → ID-Aware Tracking → Omni-Scale Embedding → Cross-Camera ReID Search → Tracklet ID Assign → Seen Entity Update → Seen Entities → Cross-Camera ReID Search



ID-aware tracking dramatically improves tracking MOTA and high FPS

CenterNet (detection) +
Kalman (trajectory) +
Embedding Distance (identity)

id	camera	exemplars	m
013	c41	[...]	0.8
013	c44	[...]	0.9

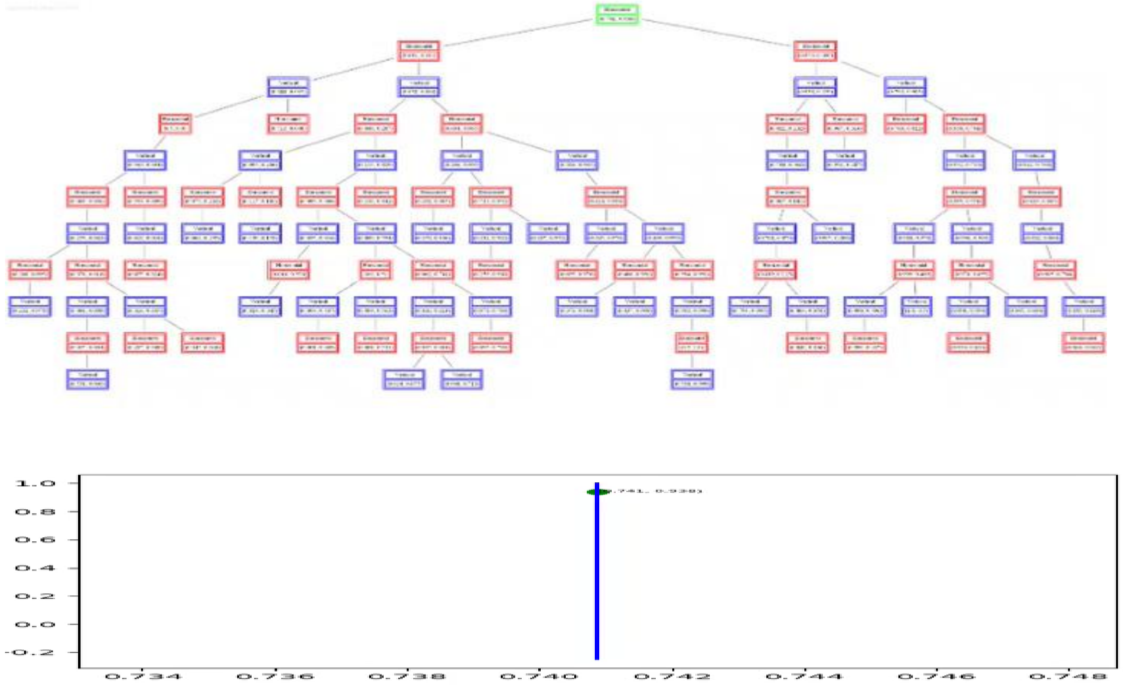
Detect **new entities**
Add **new exemplars** for known entities

[1] Zhou, Kaiyang and Yang, Yongxin and Cavallaro, Andrea and Xiang, Tao, "Omni-Scale Feature Learning for Person Re-Identification", in The IEEE International Conference on Computer Vision (ICCV), 2019

ReID Search – Efficient Vector Search

KD-Tree Search – Efficient large database vector search

Partition tree at N level using (N mod k)th vector dimension

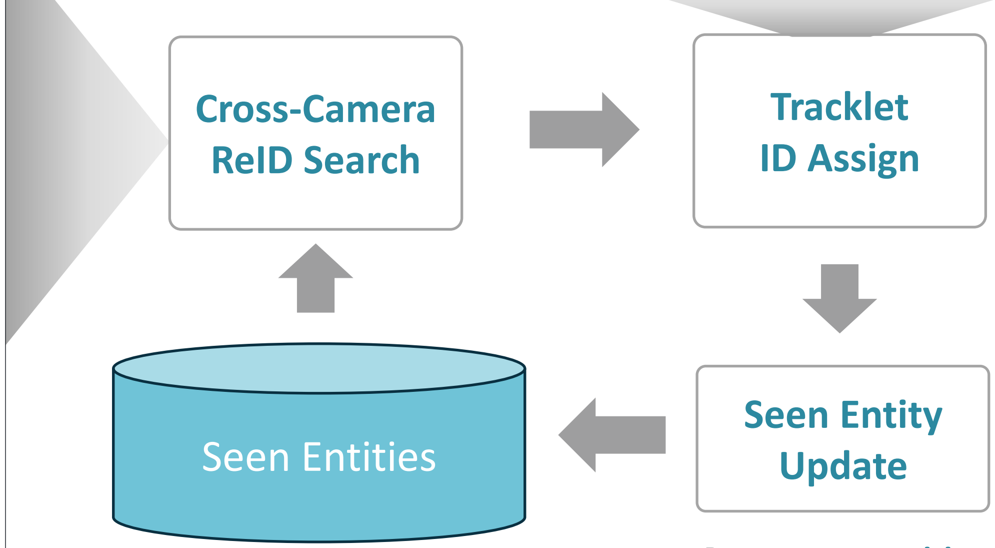


from DataScienceCentral

Tracklet ID - Assign to entire track

(smooth per-frame noisy recognition)

t	2	3	4	5	6	7
id	014	014	014	012	014	012
m	0.8	0.9	0.9	0.7	0.5	0.8



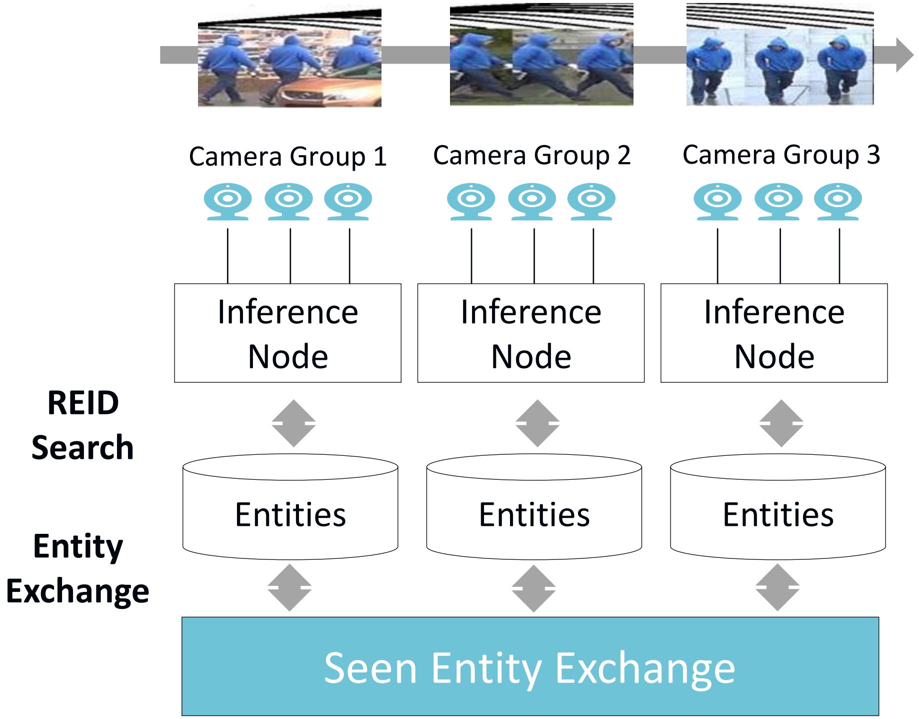
id	camera	exemplars	m
013	c41	[...]	0.8
013	c44	[...]	0.9

Detect new entities
Add new exemplars for
known entities

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

ReID Search Optimization

- Tracking requires **New IDs** to **propagate in real-time** across camera groups.
- **Multiple approaches** to balance **Recall** vs. **Compute Cost** vs. **Complexity**.
- Near-exchange provides best balance for community



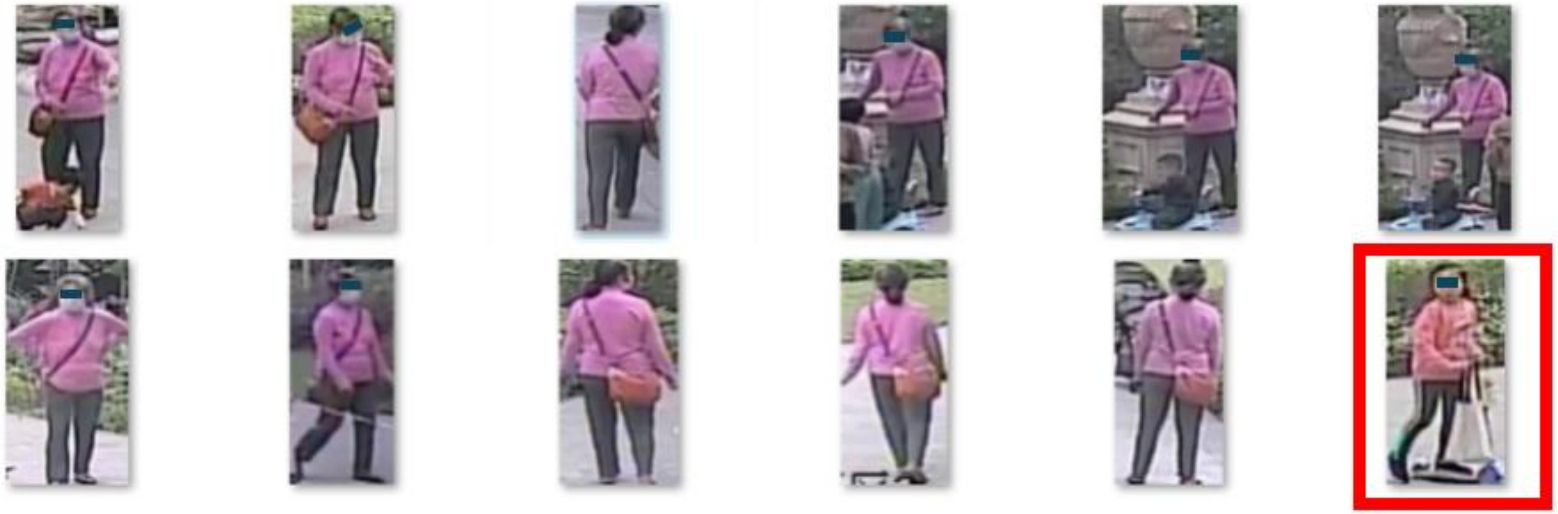
Approach	Max IDs	REID Search	Seen Entity Exchange	Latency	Recall	Complexity
Node-Inner Don't exchange IDs	N	Linear Scan (MatMult)	None	Real-Time	Poor (mitigate by heavy nodes more cameras/node)	Trivial
Global Exchange All nodes get IDs	NxK	ANN Search (NxK large)	Periodic ANN Build	>30s	Good but with latency	Complex
Near Exchange Only physically near nodes get IDs	NxV	Linear Scan (MatMult)	Real-time	Real-Time	Good	Near-Trivial
Global + Near Best of both	NxK	ANN + Linear Scan (delta)	Periodic ANN Build + Real-time (delta)	Real-Time	Best	More Complex

N = Expected Max # People / Node K = Number Nodes V = Propagate to V-nearest nodes (complexity vs. recall trade-off)

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Community Tracking Challenges

Similar clothing



Workers all look the same

Old analogue cameras

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Multi-Modal Identity Stamping

Face Identity Challenges

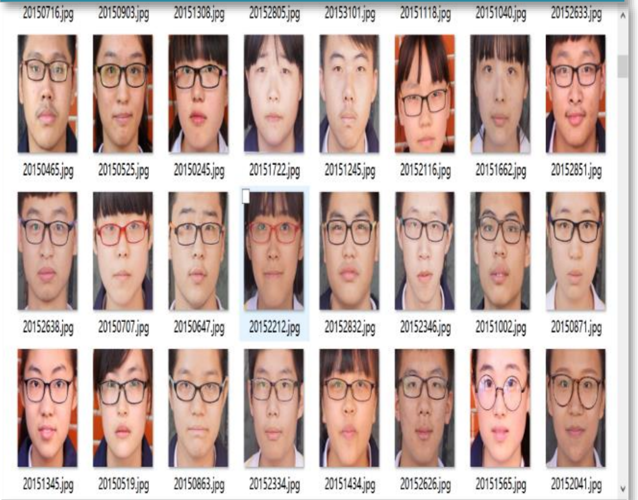
2020 embedded VISION summit

Easy

Frontal face, good lighting.

LFFW	Seedland	99.8
	Tencent	99.8
	Baidu	99.8
	Dahua	99.8

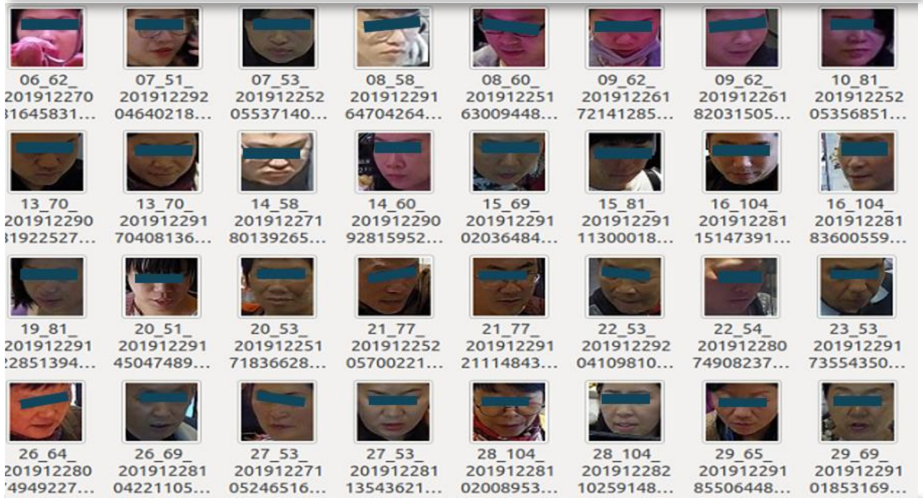
Challenging



Population bias e.g. Students

Students	Seedland	95.5@1e-6
	SenseTime '18	94.2@1e-6

Real Community Data



Lighting, angle, expression, occlusion, blur very adverse! ~92% accuracy

SEEDLAND

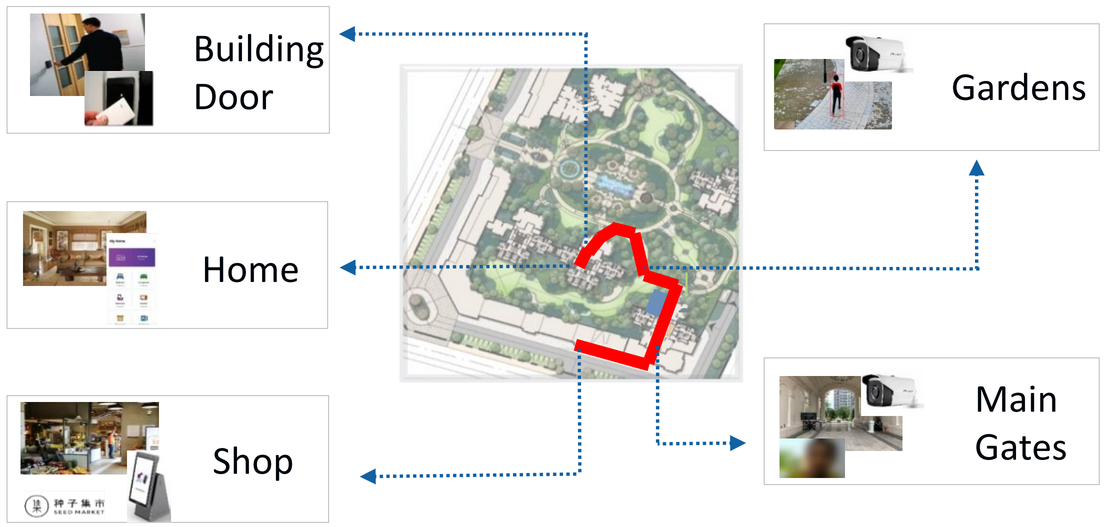
© 2020 Seedland

No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Multi-Modal Identity Stamping

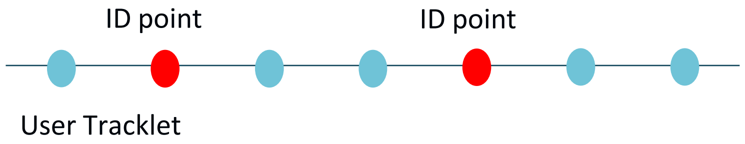
A typical user journey through the residential community has multiple touch points where user interacts with a device.

Multi-Modal identity stamping uses device interaction + visual features to convert the difficult identity task into an easier verification task.



Multi-Modal Identity Stamping

1. Use ReID to build cross-camera tracklet
2. Assign ID at key device interaction points
3. Stamp ID on tracklet after disambiguating disagreements

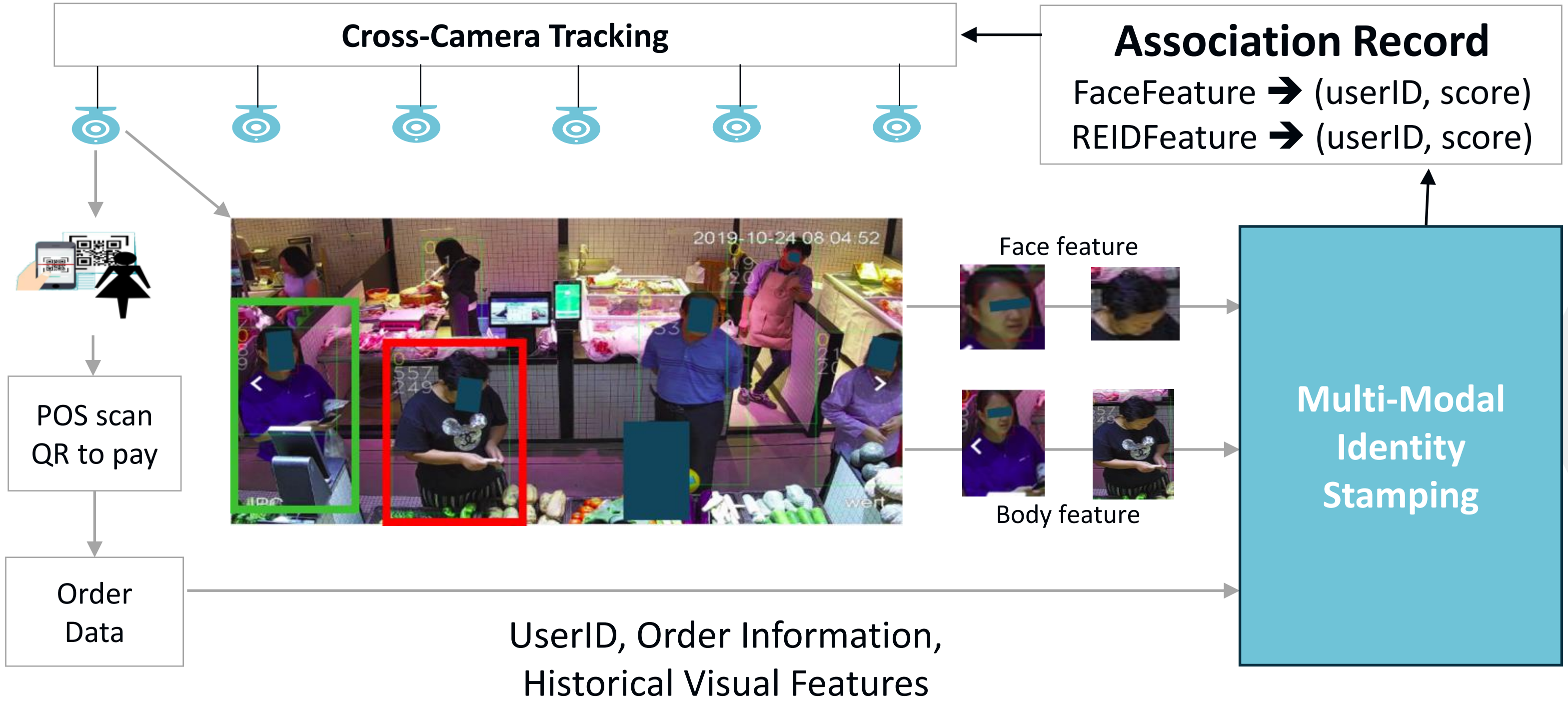


Examples of device Interactions in a user journey

- Face access systems
- Bluetooth pairing
- Swipe card access systems
- Shop POS payment

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

POS data for Multi-Modal Association


Cross-Camera Tracking
Association Record
FaceFeature → (userID, score)
REIDFeature → (userID, score)
2019-10-24 08 04:52
Face feature
Body feature
POS scan QR to pay
Order Data
Multi-Modal Identity Stamping
UserID, Order Information, Historical Visual Features

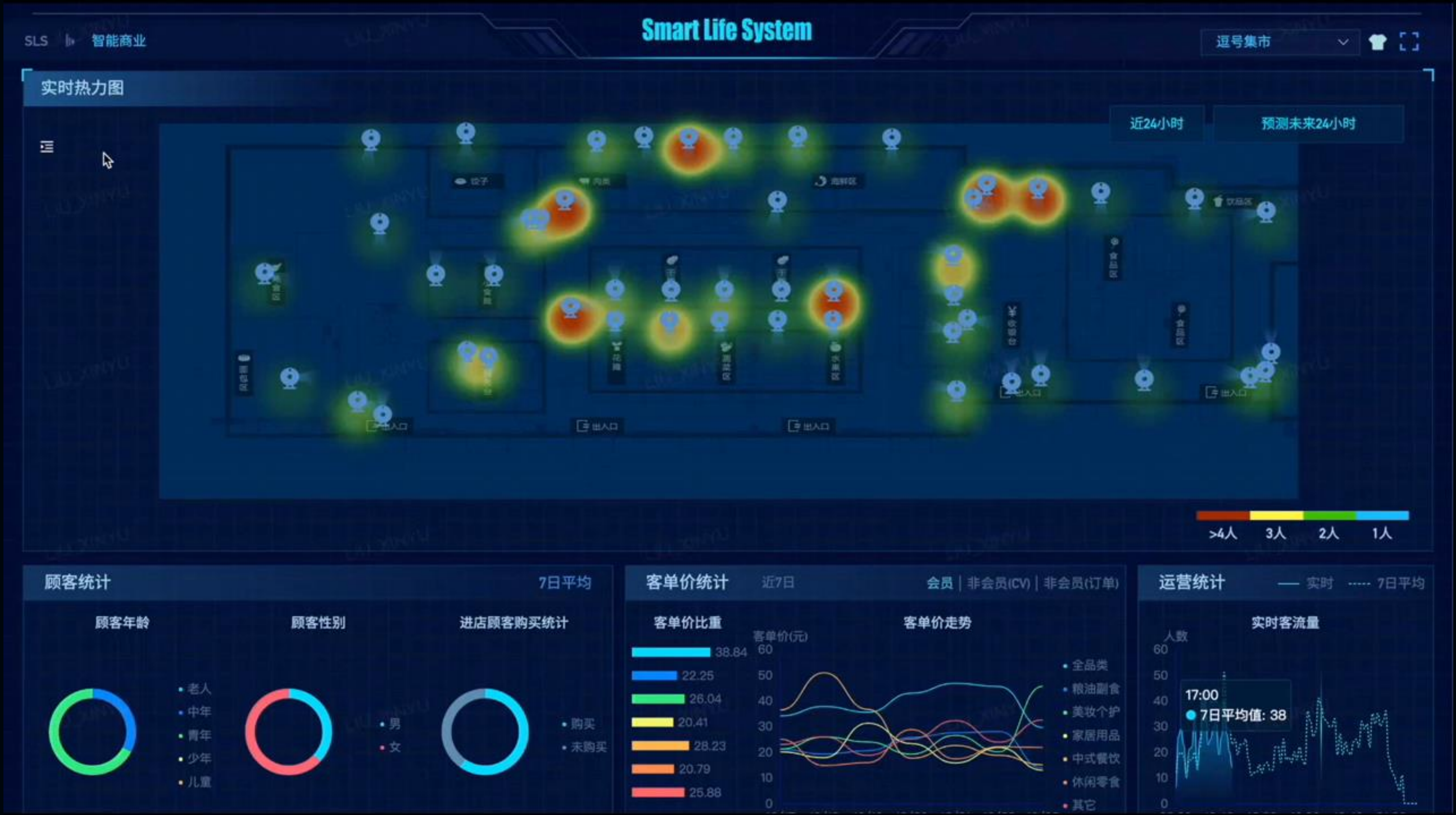
Convert hard ID task to easy Verification task

- Synchronize POS+Camera events
- Verify instead of ID face
- Store Features as ReID Exemplar for future tracking
- Generate Association Record

Increases Precision by +30% abs. over face ID. Key challenges is disambiguation – who is paying, who is waiting.

Camera Distance/Angle	Recall	Precision
Face ID only	55%	63%
Ceiling - 3m/20º	52%	90%
POS camera – 0.5m/10º	30%	93%

Retail Shop Community Tracking



SLS | 智能商业

Smart Life System

逗号集市

实时热力图

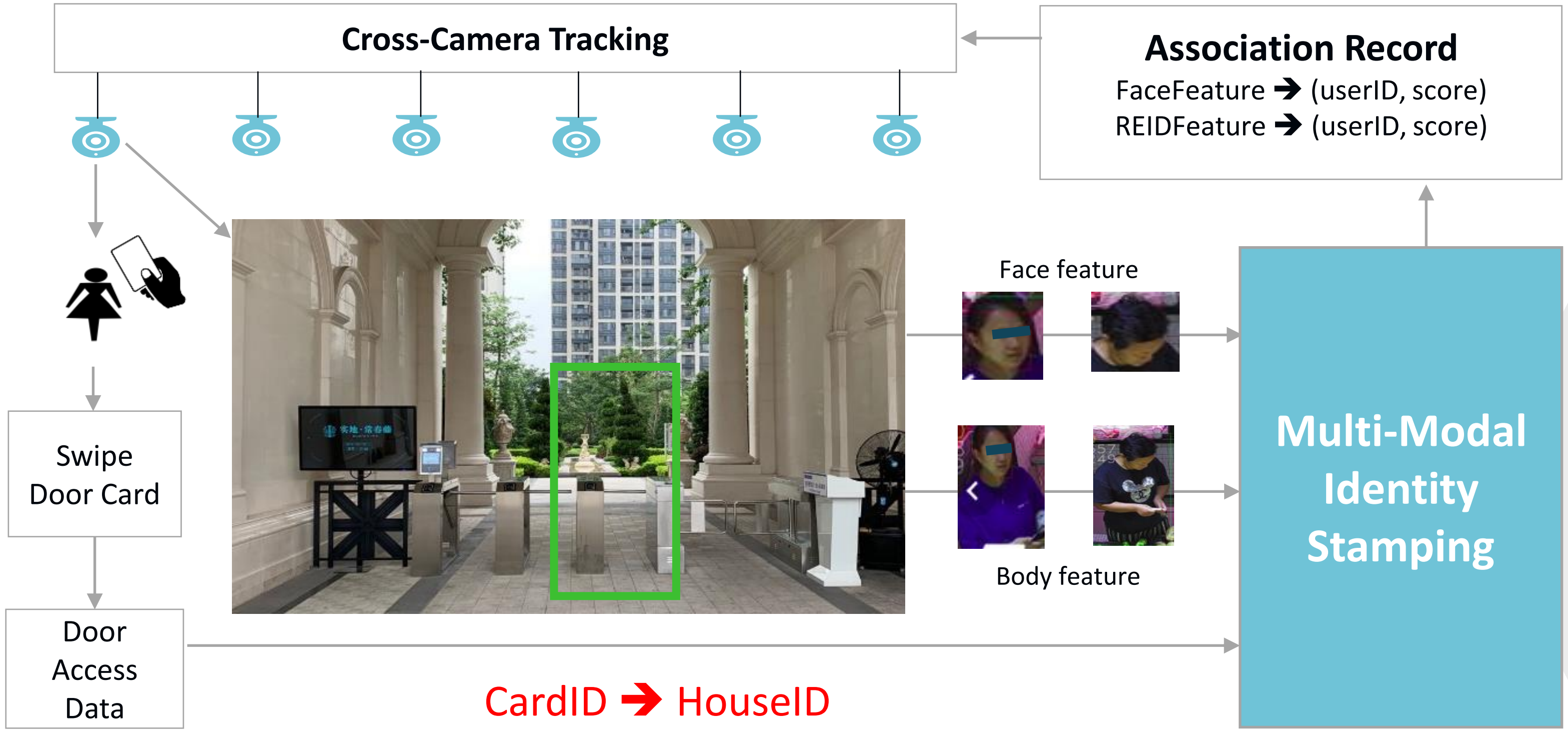
近24小时 | 预测未来24小时

顾客统计 7日平均

客单价统计 近7日 会员 | 非会员(CV) | 非会员(订单)

运营统计 — 实时 ----- 7日平均

POS data for Multi-Modal Association

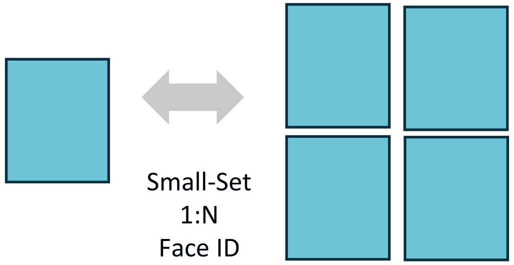


CardID ➔ HouseID

Card is not unique to a person, people may give cards to family members, visitors, ….

Propagate User ID from Door Access to Visual Features

- Accumulate visual features that historically co-occur with card ID
- 1:N identification task, N is very small (<10)

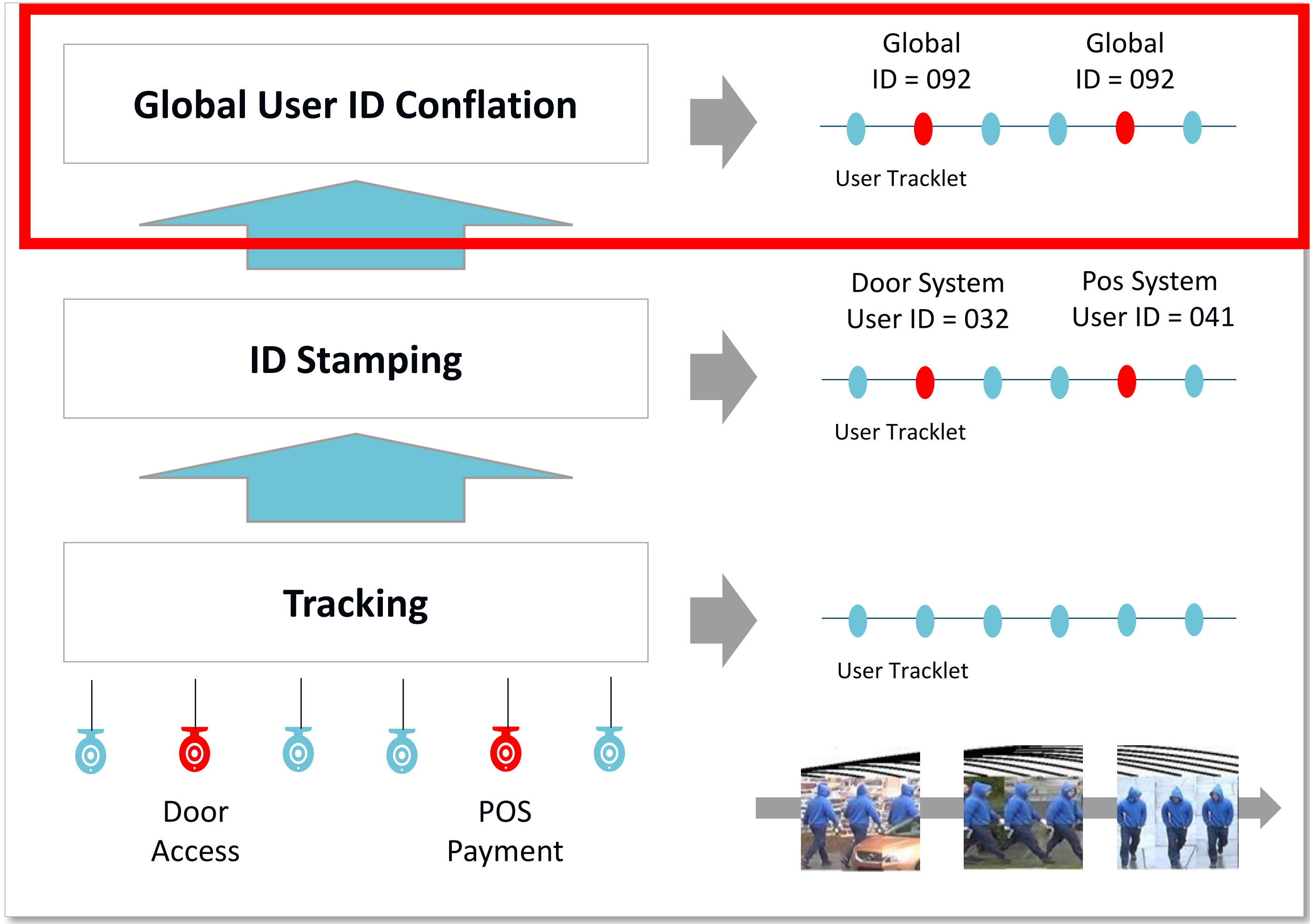


© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Global User ID Conflation

SEEDLAND

Cross-Day Stable User ID Tracking



ID stamping User IDs are system-unique, not global.

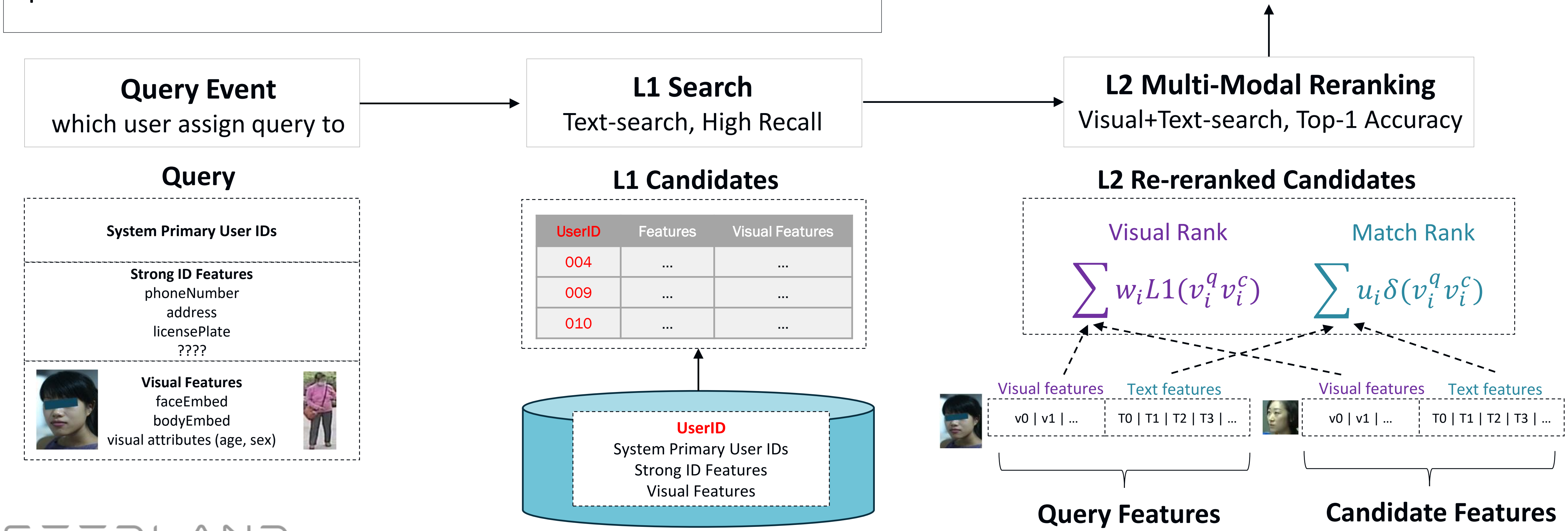
Real-Time User ID Conflation **probabilistically associates** User IDs into a **global User ID** space that is **stable and unique** per user.

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Real-Time User ID Conflation

Retrieval-based conflation does **real-time assignment** of a **Multi-Modal Query** event (text+visual) to a known set of stable **Global User IDs**. Multi-modal **increases recall by 20%** at same precision.

Choose Top-1 UserID by Rank



Query Event
which user assign query to

L1 Search
Text-search, High Recall

L2 Multi-Modal Reranking
Visual+Text-search, Top-1 Accuracy

Query

System Primary User IDs

Strong ID Features
phoneNumber
address
licensePlate
????

Visual Features
faceEmbed
bodyEmbed
visual attributes (age, sex)

L1 Candidates

UserID	Features	Visual Features
004	...	...
009	...	...
010	...	...

UserID
System Primary User IDs
Strong ID Features
Visual Features

L2 Re-reranked Candidates

Visual Rank

$$\sum w_i L1(v_i^q v_i^c)$$

Match Rank

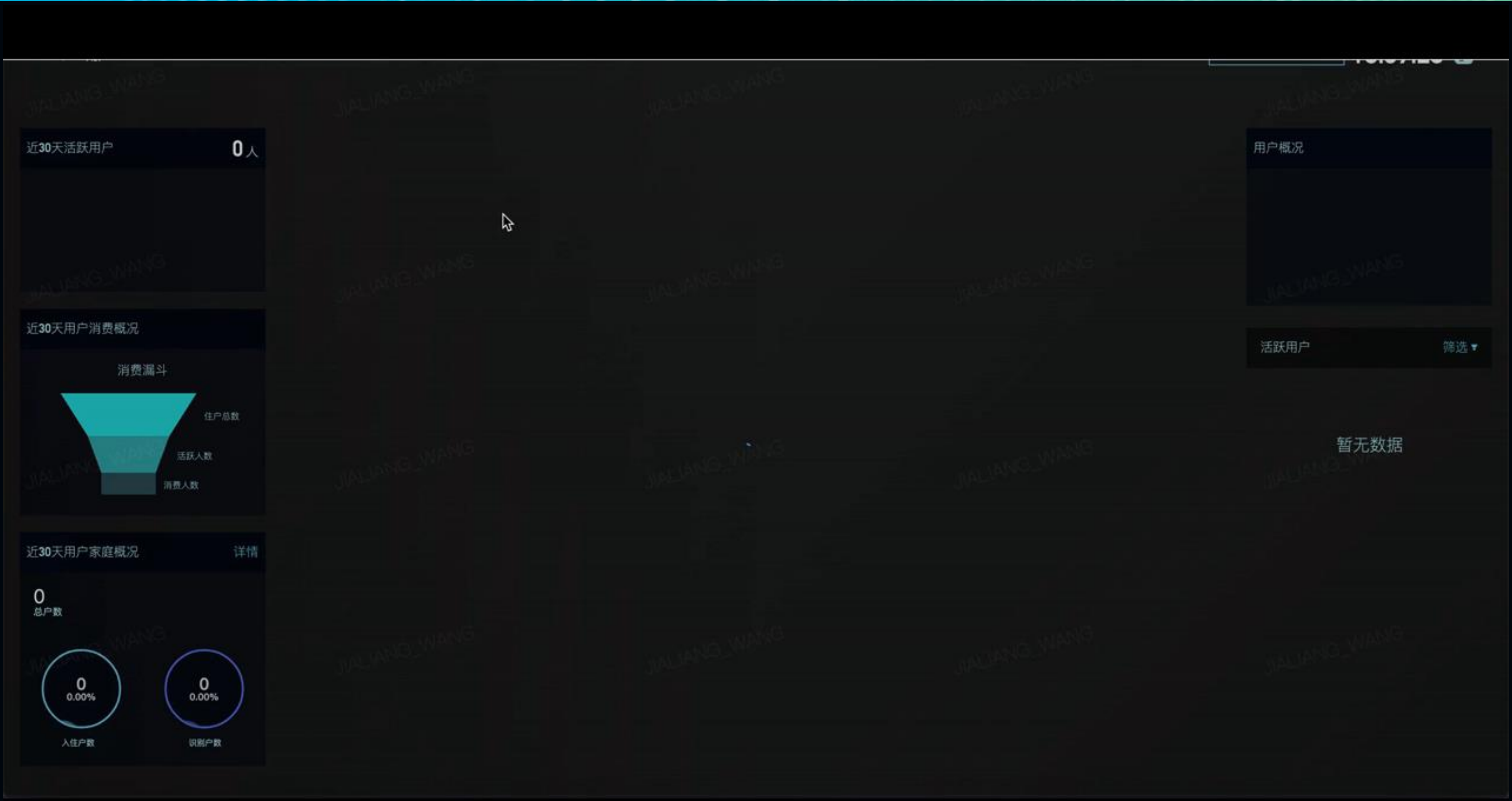
$$\sum u_i \delta(v_i^q v_i^c)$$

Query Features

Candidate Features

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland

Real-Time User Tracking and ID Conflation



近30天活跃用户 0人

近30天用户消费概况

消费漏斗

近30天用户家庭概况 详情

0
总户数

0
0.00%
入住户数

0
0.00%
识别户数

用户概况

活跃用户 筛选 ▾

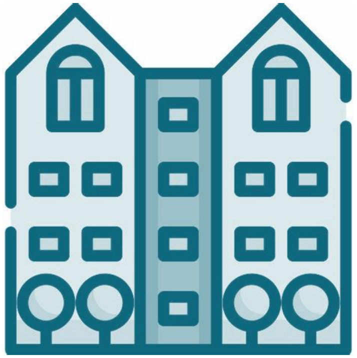
暂无数据

JIALIANG_WANG

Conclusions

SEEDLAND

Conclusions



Residential communities have many opportunities to **transform** community, lifestyle and retail **with AIOT**.

Multi-modal tracking is a vehicle for intelligent user experiences with **significant benefits over vision-only**.

Privacy and **true value** for residents must be a **first-class citizen** when playing in the tracking space.

© 2020 Seedland
No part of this presentation may be reproduced in whole or in part in any manner without the permission of Seedland