CROWD-SENSE

Crowd-Sense — Methodology and accuracy

What we measure, how we measure it, and why we publish the limits

For customers who want a defensible answer to the question "where does this number come from?" Available to attach to internal reports, council cabinet papers, and DPIA submissions. Document version 2026-04-25.

Why a methodology statement matters in 2026

In January 2026 Oxfordshire County Council publicly withdrew the November 2025 Oxford congestion-charge footfall figures after a mid-2025 dataset rebalance from its supplier Huq Industries produced anomalies the council could not defend. The supplier had not changed its methodology in any way the council was told about; the supplier had reweighted its panel.

The lesson generalises. Any council, BID, retailer, or workplace that buys footfall data without understanding — and being able to publish — the methodology behind the numbers is one supplier-side change away from the same press cycle. Auditable methodology has stopped being a nice-to-have. Crowd-Sense publishes ours and welcomes scrutiny of it.

What Crowd-Sense measures

Crowd-Sense produces **counts of unique Wi-Fi devices** observed by a sensor in a defined time window, binned by received signal strength into five distance brackets. We measure:

- **Device density** — how many devices are present in each distance bracket at each time window
- **Trend** — how density changes hour-by-hour, day-by-day, week-by-week, season-by-season
- **Comparative volume** — how density at one location or time compares to another

We do not measure:

- Absolute unique-visitor counts, especially over windows longer than a few minutes
- Individual return-visit behaviour
- Trajectories or movement paths attributable to an individual
- Demographic attributes (age, gender, segment)
- Identity of any specific person or device

These exclusions are architectural: device identifiers are discarded on the sensor before any data leaves it (see `DPO_TECHNICAL_NOTE.pdf`). The wire payload contains nothing that could support an identity-level claim.

The MAC randomisation overcount — full disclosure

Modern smartphones (iOS 14 and later, Android 10 and later, both shipping since 2019–2020) randomise the MAC addresses they broadcast in pre-association probe requests, by default. A single phone passing a sensor for an extended period is observed under multiple different randomised MAC addresses as the operating system rotates them.

This rotation is not a bug; it is a privacy feature deliberately designed by operating-system vendors to defeat passive tracking. As a sensor that **respects** that privacy feature (rather than attempting to defeat it), Crowd-Sense systematically overcounts unique-device counts at longer time windows. The empirical overcount factors:

Time window	Typical overcount factor	Operationally meaningful for
1 minute	**1.2× to 2×**	Live event-time decisions; queue management; rapid response
5 minutes	**1.5× to 3×**	Short-term trend; staff allocation
15 minutes	**2× to 5×**	Operational reporting; cleaning triggers; HVAC trigger decisions
30 minutes	**3× to 7×**	Hourly trend; capacity management
60 minutes	**4× to 10×**	Long-window density and volume reporting

These factors are venue-dependent, and depend on dwell-time distribution, device-mix, signal environment, and ambient Wi-Fi activity. They are not knowable at the sensor without additional ground-truth data.

How customers should use these factors

The single most common mistake is to use long-window unique-device counts as if they were unique-visitor counts. They are not. For a typical UK high street with average mid-day dwell, a "60 unique devices in the 8–15 m bracket between 12:00 and 13:00" reading from a single sensor probably corresponds to 6–15 actual humans at any given moment.

For density and trend (the use case Crowd-Sense is designed for), the overcount factor is approximately constant for a given venue — so relative comparisons (busier than yesterday? busier than last hour? busier than the same Saturday last year?) are reliable, even though absolute device counts inflate the headcount.

For absolute volume (if you genuinely need "how many distinct humans walked past this point?"), use a 1-minute window where the overcount factor is small, then sum windows over the period of interest with a known multiplier discount. Or use a complementary deterministic counter (entry-gate counter, ticket sales, POS) to calibrate.

For event-time decisions ("is this venue busy right now?"), use the 1-minute window directly. The overcount is small enough that operational decisions are unaffected.

From 2027, Crowd-Sense subscriptions will include venue calibration: an ML model that learns the local overcount factor against a customer-supplied ground-truth signal (POS, ticket sales, manual counts) and converts device counts into estimated effective humans. Until that capability ships, the figures above are the published rule-of-thumb.

What about vendors claiming 95%+ accuracy?

Any supplier offering Wi-Fi-derived "unique visitors over an hour" with 95%+ accuracy in 2026 is doing one of three things:

1. **Re-identifying randomised MACs.** Techniques exist that link probes from the same device across MAC rotations — Information Element entropy fingerprinting, sequence-number tracking, timing-pattern correlation, RSSI-trajectory matching. These work, with varying success, but they actively reverse the privacy feature the user's operating system applied. Under UK GDPR and the WP216 anonymisation tests, this is legally contestable. The CNIL refused JCDecaux's 2015 La Défense pilot on this exact ground; the Conseil d'État upheld the refusal in 2017.
2. **Mixing in non-Wi-Fi signals.** Captive-portal associations (where the customer logs in to free venue Wi-Fi), BLE beacons paired with consenting apps, or in-venue cameras can produce more accurate counts. These are not Wi-Fi probe analytics; they are different products being marketed under the same banner.
3. **Overclaiming.** The 95% number is unverifiable without an independent ground truth. Most vendors do not publish their measurement methodology and decline customer requests to do so.

We invite customers to ask any competing vendor which of the three applies to their accuracy claim, and to ask for the published methodology behind it. The answer is informative regardless of which of the three is given.

What we ask customers to do with this information

When citing Crowd-Sense data in published reports, cabinet papers, BID annual reviews, or comparable documents, we ask that the methodology be summarised honestly. Suggested form of words:

> *"Footfall data sourced from Crowd-Sense passive Wi-Fi sensors. The sensors count unique Wi-Fi devices broadcasting in proximity to the sensor, and report aggregate counts only — device identifiers are discarded on the sensor and never transmitted. Modern mobile devices randomise their broadcast MAC addresses, which causes a known overcount of unique humans at longer time windows (approximately 1.2–2× at 1-minute windows; 2–5× at 15-minute windows; 4–10× at 60-minute windows). These figures should be read as device-density and trend indicators, not absolute unique-visitor counts."*

That paragraph fits in a cabinet paper footnote. We provide it because we would rather customers cite it correctly than have the methodology become a vulnerability later.

Document scope

This methodology applies to Crowd-Sense passive Wi-Fi probe sensors as currently designed. It does not cover camera, thermal, depth, BLE-beacon, or SDK-aggregated location-intelligence products — those have different accuracy and privacy characteristics, and should be evaluated on their own published methodology.

For the technical architecture supporting these claims, see the companion documents:

- *Technical privacy note for Data Protection Officers*
- *Technical specification* (available under NDA)
- *DPIA template* (available on request)

Crowd-Sense is a product of **Visual Solutions UK Limited** (Companies House #02955160). For questions or corrections, contact Jeff Burton, Director (jeff.burton@crowd-sense.com).