

Article

Ahead of the New Panini Mania Wave: How Much Does It Cost to Complete the Entire FIFA World Cup 2026 Sticker Album?

Diego Šumberac and Petar Kolar * 

VERN' University, Palmotićeve ulica 82/1, HR-10000 Zagreb, Croatia; diego.sumberac@proton.me

* Correspondence: petar.kolar@vern.hr

Abstract

The FIFA World Cup 2026 will be accompanied by the largest official sticker album released so far. Due to the expansion of the tournament from 32 to 48 national teams, the album size has increased to 980 stickers, while the number of stickers per pack has increased from 5 to 7. Such an expansion is expected to increase the total completion cost. However, this increase is not linear, because the collection process becomes dominated by duplicate accumulation in its final stage. The aim of this paper is to estimate, a priori, the expected number of packs, total number of stickers, and financial cost required to complete the album under a uniform random-acquisition baseline. The process is modeled using an exact finite-state recursive formulation for multi-item sampling, while Monte Carlo simulations based on 10,000 independent trials are used to characterize the distribution and variability of completion outcomes. Under the uniform random-acquisition baseline, the exact recursive expectation is about 1042 packs, while the Monte Carlo mean is about 1044 packs. This corresponds to approximately 7311 purchased stickers in the Monte Carlo estimate, of which only about 13% contribute directly to completion, while 87% are duplicates. The engineering relevance of the acquisition structure is examined through a real LoRaWAN dataset case study, where address-level first coverage, repeated device observations, and packet de-duplication show analogous long-tail acquisition behavior in an IoT data-acquisition system.

Keywords: cost analysis; Coupon collector's problem; data acquisition; Monte Carlo simulation; random sampling; stochastic processes

1. Introduction

Official sticker albums accompanying the FIFA World Cup have been published since the FIFA World Cup 1970, when the final tournament involved 16 national teams and the corresponding album contained 288 stickers [1]. Since then, the number of participating teams has increased gradually: first from 16 to 24 in 1982, then from 24 to 32 in 1998, and now from 32 to 48 in the FIFA World Cup 2026 [2,3]. Expectedly, the corresponding sticker album also became larger over time. In the editions associated with the FIFA World Cup 2018 and the FIFA World Cup 2022, the number of stickers reached 670 [1]. In the FIFA World Cup 2026 edition, it has increased further to 980 [4].

The number of stickers per pack has also changed, from 5 to 7. At first sight, this may suggest that the album has simply grown and that the completion cost should increase accordingly. However, the main problem is not the album size itself. The main problem is the stochastic nature of the acquisition process. As the collection progresses, the probability of obtaining a previously unseen sticker decreases, whereas the number of duplicates



Academic Editor: Francesco Zirilli

Received: 21 June 2026

Revised: 16 July 2026

Accepted: 21 July 2026

Published: 24 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

increases. Because of that, the total cost of completing the album grows much faster than a simple linear estimate would suggest.

This phenomenon can be described using the coupon collector's problem [5], a classical probabilistic model that describes how many random samples are required in order to obtain all elements from a finite set. The problem has been discussed before in the context of sticker collections, including analytical and heuristic considerations of cost and strategy [6]. The present work does not aim to introduce a new class of stochastic processes. Instead, its contribution is to provide a reproducible baseline analysis of a large multi-item sticker collection system by combining an exact finite-state expectation, Monte Carlo distributional results, sensitivity testing, and practical cost interpretation in a single consistent framework.

In the present work, the classical coupon collector's problem is used only as a reference case and as a starting point for interpretation. The actual acquisition process considered here is different from the classical single-item formulation, because one purchase step yields several mutually distinct stickers. For that reason, the theoretical treatment is based on a finite-state stochastic formulation of the multi-item acquisition process, while the classical result is retained only as a useful benchmark.

A preliminary quantitative analysis of the coupon collector's problem applied to the FIFA World Cup 2022 sticker album was previously reported by the authors in [7]. That study considered a smaller system with 670 stickers and 5 stickers per pack. The present paper extends that approach in two main directions. First, it considers a substantially larger system corresponding to the FIFA World Cup 2026 album. Second, and more importantly, it provides a more explicit and consistent mathematical treatment of the multi-item acquisition process, together with a detailed interpretation of its impact on the total completion cost and its variability.

The aim of this paper is to estimate, *a priori*, the number of packs required to complete the FIFA World Cup 2026 sticker album if the collector relies only on random pack purchases. To that end, the collection process is modeled as a coupon collector system with multi-item sampling. Based on the developed model, the expected number of required packs is estimated analytically and then verified by Monte Carlo simulations. In addition to the average value, the spread of the completion process is also analyzed in order to estimate realistic favorable and unfavorable completion scenarios.

This uniform random-acquisition scenario is intentionally treated as a baseline reference case, not as a description of an optimal or typical real-world collecting strategy. It should also not be interpreted as a strict upper bound, because non-uniform sticker availability could increase the completion time, while exchange, targeted purchase, and secondary-market acquisition could reduce the actual cost incurred by collectors. Practical strategies involving such mechanisms are discussed separately in the cost analysis section.

In order to address the engineering relevance of the same acquisition structure, the paper also analyzes an existing real-world LoRaWAN dataset. The LoRaWAN at the Edge Dataset (LoED) contains gateway-level LoRaWAN packet traces collected in an urban deployment and has been used in previous work for LoRaWAN traffic analysis, including the study of packet de-duplication [8,9]. In the present paper, this dataset is not used to validate the sticker-album model as a LoRaWAN performance model. Instead, it is used as a realistic engineering case study in which first coverage, repeated observations, non-uniform acquisition, and redundancy can be measured from real data.

Accordingly, the novelty of the paper is primarily application-oriented and interpretative, rather than methodological in the sense of introducing a new probability model. The main contributions of this paper are as follows:

1. An application of a finite-state stochastic formulation to the FIFA World Cup 2026 sticker collection process with multi-item sampling without replacement within each pack,
2. A comparison between the classical coupon collector benchmark, the mean-field approximation, the exact recursive expectation, and Monte Carlo simulation results,
3. A detailed numerical characterization of the completion process, including the expected number of packs, the derived duplicate fraction, empirical percentiles, and variability across simulations,
4. A baseline cost estimate for the purely random pack-purchase scenario, together with an explicit discussion of its limitations with respect to trading, targeted purchase, and other practical collecting strategies,
5. A real LoRaWAN dataset case study based on the LoED packet traces, used to examine address-level first coverage, repeated observations, and redundancy in an engineering data-acquisition context.

Thus, the main goal of this paper is not only to estimate the expected completion cost of the FIFA World Cup 2026 sticker album but also to explain why that cost becomes so high.

2. Theoretical Model

The process of completing a sticker album can be modeled as a stochastic acquisition process in which a finite set of n distinct items is collected through repeated random sampling [10]. In the considered case, the total number of distinct stickers is $n = 980$, and each sampling step corresponds to the purchase of one pack containing $k = 7$ stickers.

The following assumptions are adopted:

- All stickers are equally likely to appear;
- Stickers are sampled independently across different packs;
- No duplicate sticker appears within a single pack.

Under these assumptions, the considered process can be viewed as an extension [5] of the classical coupon collector's problem [10,11], in which multiple distinct items are obtained in one trial.

The following subsection first recalls the classical coupon collector result in order to define the reference case. After that, the model is extended to the actual pack-based acquisition process by using transition probabilities between collection states. In this way, the multi-item case is not treated only as a direct division of the classical result by the pack size but as a separate finite-state stochastic process.

2.1. Classical Coupon Collector's Problem

In the classical formulation, each trial yields exactly one randomly selected item from a set of n possible items. Hence, only one element is acquired per acquisition step, independently and uniformly at random.

The expected number of trials required to collect all n distinct items is given by

$$E(T_1) = nH_n = n \sum_{i=1}^n \frac{1}{i} \approx n \ln n + \gamma n, \quad (1)$$

where T_1 denotes the number of trials required to collect all n distinct items in the case $k = 1$ [10,12]. The quantity H_n is the n -th harmonic number, defined as

$$H_n = \sum_{i=1}^n \frac{1}{i}. \quad (2)$$

The harmonic number reflects the cumulative difficulty of obtaining new items during the collection process. At the beginning of the process, the probability of obtaining a previously unseen item is high. Later, as the number of already collected items increases, the probability of obtaining one of the remaining missing items decreases. Since H_n grows logarithmically [12], it can be approximated as

$$H_n \approx \ln n + \gamma, \quad (3)$$

where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant, which arises in the asymptotic relation between the harmonic series and the natural logarithm.

This result shows that the expected number of required samples does not grow linearly with the collection size. The reason is the increasing difficulty of obtaining the last few missing items. This is the well-known “long tail” effect of the coupon collector’s problem [11].

2.2. Exact Finite-State Formulation for Multi-Item Sampling

In the considered album, each acquisition step yields $k = 7$ distinct stickers sampled without replacement from the set of $n = 980$ possible stickers. Therefore, the process is not identical to the classical coupon collector’s problem with k independent single-sticker trials. The stickers inside one pack are mutually distinct, which changes the transition probabilities between collection states.

Let m denote the number of distinct stickers already collected. The number of missing stickers is then $n - m$. If a new pack contains exactly j previously unseen stickers, then the collection state changes from m to $m + j$. Under the assumptions of uniform sampling and no duplicate stickers within a pack, the probability of obtaining exactly j new stickers in one pack is given by the hypergeometric probability

$$P_{m,j} = \frac{\binom{n-m}{j} \binom{m}{k-j}}{\binom{n}{k}}, \quad (4)$$

where

$$\max(0, k - m) \leq j \leq \min(k, n - m). \quad (5)$$

For values of j outside this range, the corresponding transition probability is interpreted as zero. This formulation explicitly accounts for the fact that the k stickers in one pack are sampled without replacement. The term $\binom{n-m}{j}$ counts the possible choices of new stickers, while $\binom{m}{k-j}$ counts the possible choices of already collected stickers that appear as duplicates with respect to the current collection.

Let E_m denote the expected number of additional packs required to complete the album when m distinct stickers have already been collected. Clearly,

$$E_n = 0. \quad (6)$$

For $m < n$, the expectation satisfies the recursive relation

$$E_m = 1 + \sum_{j=0}^{\min(k, n-m)} P_{m,j} E_{m+j}. \quad (7)$$

When $j = 0$, the new pack contains no previously unseen stickers, and the process remains in the same state m . Therefore, Equation (7) can be rearranged as

$$E_m = \frac{1 + \sum_{j=1}^{\min(k, n-m)} P_{m,j} E_{m+j}}{1 - P_{m,0}}. \quad (8)$$

Equation (8) gives an exact finite-state recursive formulation of the expected completion time under the assumptions used in this paper. It was evaluated backwards from $m = n$ to $m = 0$. For $n = 980$ and $k = 7$, this gives

$$E_0 = 1042.36 \text{ packs}. \quad (9)$$

This value represents the exact expected number of packs for the simplified random acquisition model with uniform sticker probabilities, independent packs, and no duplicate stickers within a single pack.

2.3. Mean-Field Approximation

For comparison, it is still useful to consider a simpler mean-field approximation. At a stage characterized by m already collected stickers, the expected number of newly obtained stickers in one pack is

$$E[\Delta m] = k \cdot \frac{n-m}{n}. \quad (10)$$

This expected increment is exact as an expectation, although it does not describe the full transition distribution. If the reciprocal of this expected increment is used as an approximate estimate of the number of packs required to increase the collection from m to $m+1$, then

$$E(T_k)_{\text{MF}} \approx \sum_{m=0}^{n-1} \frac{1}{k \cdot \frac{n-m}{n}} = \frac{n}{k} H_n. \quad (11)$$

Using the asymptotic approximation $H_n \approx \ln n + \gamma$, this gives

$$E(T_k)_{\text{MF}} \approx \frac{n}{k} (\ln n + \gamma). \quad (12)$$

For $n = 980$ and $k = 7$, the mean-field approximation gives

$$E(T_k)_{\text{MF}} \approx \frac{980}{7} (\ln(980) + 0.5772) \approx 1045.07 \text{ packs}. \quad (13)$$

Compared with the exact recursive value $E_0 = 1042.36$ packs, the relative difference is approximately 0.26%. Therefore, the mean-field expression provides a close numerical approximation for the considered parameter values, but it is not treated here as an exact analytical result. In the revised manuscript, the exact recursive formulation is used as the primary theoretical result, while the mean-field expression is retained only as a simple benchmark and interpretation tool.

The mean-field approximation is used only for the expected value. Favorable and unfavorable completion scenarios are not derived from this approximation. Instead, they are reported from the empirical distribution obtained by Monte Carlo simulation. This keeps the interpretation of the completion range directly connected to the simulated pack-based acquisition process.

3. Simulation Methodology

In order to complement the theoretical analysis and to assess the variability of the considered acquisition process, Monte Carlo simulations were performed [13,14]. The simulations follow directly from the stochastic model introduced in Section 2. The applied implementation builds upon previously developed numerical treatments of the coupon

collector process in practical collection systems [7], but is extended here to a larger collection size, multi-item sampling, and a broader statistical analysis.

In each simulation run, an initially empty collection of $n = 980$ distinct stickers is considered. At every acquisition step, one pack containing $k = 7$ stickers is generated by uniform random sampling from the complete set of stickers, without replacement within the pack itself. The process is repeated until all n distinct stickers have been collected.

During each simulation run, the following quantities are recorded:

- The total number of packs required to complete the album;
- The total number of purchased stickers;
- The number of duplicate stickers derived from the total number of purchased stickers after completion.

A total of 10,000 independent simulation runs were performed in order to ensure statistical reliability of the obtained results [14].

3.1. Simulation Procedure

The simulation procedure can be summarized as follows:

1. Initialize an empty set representing the currently collected stickers,
2. Repeat the acquisition process until all n stickers have been collected:
 - Generate one pack of k distinct stickers sampled uniformly at random,
 - Add all newly obtained stickers to the collection,
 - Increment the pack counter,
3. Record the total number of required packs and derive the corresponding number of duplicate stickers.

3.2. Averaging and Statistical Analysis

Besides the total number of packs required for album completion, the evolution of the acquisition process was also analyzed by observing the average number of distinct stickers collected as a function of the number of purchased packs.

More precisely, for each pack index, t , the average number of collected stickers was computed over all simulation runs [14]. In that way, an average acquisition curve was obtained, which describes the typical progress of the collection process.

As a reference, an idealized linear scenario was also considered. In that hypothetical case, every newly purchased sticker is assumed to be different from all previously purchased ones, so the number of collected stickers grows simply as $k \cdot t$. This reference curve serves only as a lower-bound benchmark because, in any realistic random acquisition process, duplicates inevitably appear.

3.3. Output Metrics

The simulation results are evaluated using the following quantities:

1. The expected number of packs required for completion;
2. The distribution of the required number of packs;
3. The median, standard deviation, standard error, confidence interval, and empirical percentiles of the required number of packs;
4. The convergence of the estimated mean value as the number of Monte Carlo trials increases;
5. The minimum and maximum observed completion values;
6. The derived number and fraction of duplicate stickers;
7. The average evolution of the number of collected stickers.

These quantities provide insight not only into the expected behavior of the considered system but also into its variability.

3.4. Implementation Details

The simulations were implemented in Python 3.13.5 using standard scientific libraries, including NumPy 2.3.5 for numerical processing and Matplotlib 3.10.8 for visualization. The implementation follows directly from the introduced stochastic model and allows the efficient execution of a large number of independent trials. In order to make the numerical results exactly reproducible, the pseudo-random number generator was initialized using a fixed seed equal to 2026. The main 10,000-run baseline simulation used seed 2026, whereas the independent-batch verification described in the Results section used 10 separate 10,000-run batches initialized with different random seeds.

For reproducibility reasons, the implementation is publicly available through an open-source GitHub repository [15].

4. Results and Discussion

The results obtained by Monte Carlo simulations [14] provide quantitative insight into both the expected behavior of the sticker collection process and its stochastic variability.

4.1. Average Collection Dynamics

The average progress of the collection process is shown in Figure 1, where the number of collected stickers is plotted as a function of the number of purchased packs. The displayed curve represents the average over 10,000 independent simulation runs, while the linear reference corresponds to the idealized duplicate-free case.

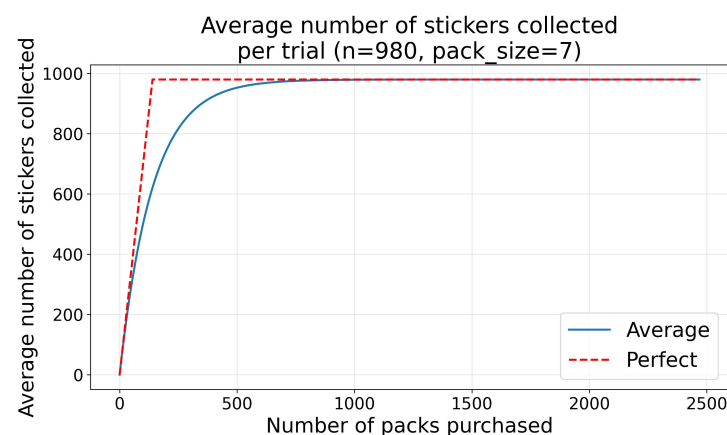


Figure 1. Average number of collected stickers as a function of the number of purchased packs, compared to the idealized linear case without duplicates.

At the beginning of the process, the average acquisition curve follows the linear reference relatively closely. This means that, in the early stage of the collection, the probability of obtaining new stickers is high, and the acquisition process is efficient. As the number of already collected stickers increases, the curve departs gradually from the linear reference. This departure becomes especially pronounced near the completion of the album, where the probability of obtaining previously unseen stickers drops significantly.

Such behavior is fully consistent with the theoretical model presented in Section 2. More importantly, it confirms that the main inefficiency of the considered acquisition process is concentrated in its final stage, i.e., in the long-tail region of the coupon collector's problem [11].

4.2. Distribution of Required Packs

The distribution of the total number of packs required to complete the album is shown in Figure 2. This figure represents the Monte Carlo distribution of completion outcomes. The Monte Carlo mean is about 1044 packs, which is close to the exact recursive expectation of about 1042 packs for the same baseline model.

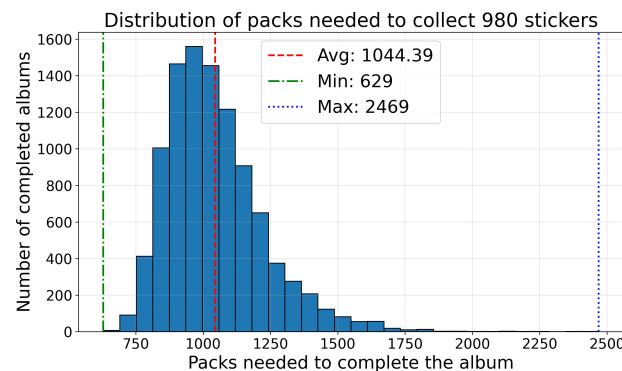


Figure 2. Distribution of the total number of packs required to complete the album over 10,000 Monte Carlo simulations.

The histogram is right-skewed, which is typical for coupon-collector-type completion processes. Although many realizations are concentrated around the mean, the right tail shows that substantially less favorable outcomes are still possible. Therefore, the empirical percentiles reported below are used to describe favorable and unfavorable completion scenarios.

4.3. Statistical Summary and Convergence of Simulations

Before discussing the simulation distribution in detail, Table 1 summarizes the three numerical quantities used in the manuscript. The exact recursive expectation is the theoretical expected value of the baseline finite-state model. The Monte Carlo mean is a finite-sample estimate obtained from 10,000 simulated realizations of the same process. The mean-field approximation is retained only as a simple benchmark approximation.

Table 1. Separation of the exact recursive result, Monte Carlo estimate, and mean-field approximation for the uniform random-acquisition baseline.

Quantity	Value	Interpretation
Exact recursive expectation	1042.36 packs	Theoretical expectation of the finite-state baseline model
Monte Carlo mean	1044.39 packs	Finite-sample estimate from 10,000 simulation runs
Mean-field approximation	1045.07 packs	Simple benchmark approximation

In order to describe the Monte Carlo results more completely, the simulated distribution of the required number of packs was evaluated using standard descriptive statistics. These values refer only to the Monte Carlo sample and are summarized in Table 2.

The mean value is close to the exact recursive expectation derived in Section 2. The median is lower than the mean, which is consistent with the right-skewed shape of the distribution. The empirical 5th and 95th percentiles are used to describe favorable and unfavorable completion scenarios. Therefore, these scenarios are obtained directly from the Monte Carlo distribution, rather than from an additional analytical approximation.

The convergence of the Monte Carlo estimate was evaluated by computing the running mean after different numbers of independent simulation runs. The corresponding values are shown in Table 3.

Table 2. Summary statistics of the Monte Carlo simulation results for the purely random acquisition scenario.

Quantity	Value
Number of simulation runs	10,000
Mean number of packs	1044.39
Median number of packs	1014
Standard deviation	181.02
Standard error of the mean	1.81
95% confidence interval of the mean	[1040.85, 1047.94]
5th percentile	811
95th percentile	1387
Minimum observed value	629
Maximum observed value	2469

Table 3. Convergence of the estimated mean number of required packs with respect to the number of Monte Carlo simulation runs.

Number of Runs	Estimated Mean	95% Confidence Interval
100	1030.47	[995.06, 1065.88]
500	1032.67	[1016.58, 1048.76]
1000	1035.67	[1024.66, 1046.67]
2500	1047.65	[1040.28, 1055.02]
5000	1043.93	[1038.87, 1049.00]
7500	1044.72	[1040.59, 1048.84]
10,000	1044.39	[1040.85, 1047.94]

As expected, the confidence interval becomes narrower as the number of simulation runs increases. This indicates that the Monte Carlo estimate stabilizes as the number of trials increases.

In addition to the running-mean convergence check, an independent-batch verification was also performed. The baseline simulation was repeated using 10 independent batches, each consisting of 10,000 simulation runs and initialized with a different random seed. This provides an additional check that the estimated mean is not dependent on a single simulation sequence. The obtained batch-level results are summarized in Table 4.

Table 4. Independent-batch verification of the Monte Carlo estimate for the uniform random-acquisition baseline.

Quantity	Value
Number of independent batches	10
Runs per batch	10,000
Mean of batch means	1043.48 packs
Standard deviation of batch means	1.64 packs
Minimum batch mean	1041.18 packs
Maximum batch mean	1046.19 packs

The mean of the independent batch means is close to both the exact recursive expectation of 1042.36 packs and the main Monte Carlo mean of 1044.39 packs. The small spread between batch means indicates that the reported baseline estimate is stable with respect to different random seeds.

4.4. Duplicate Stickers as a Derived Quantity

Duplicate stickers are treated as a derived quantity rather than as an independent stochastic output, because their number follows directly from the number of purchased packs. If P denotes the number of purchased packs, then the total number of purchased stickers is kP , and the number of duplicate stickers after completion is

$$D = kP - n. \quad (14)$$

Therefore, the distribution of duplicate stickers contains the same information as the distribution of required packs shown in Figure 2. The duplicate count is consequently reported through the derived relation above and the corresponding average values.

Using the Monte Carlo mean of 1044.39 packs, the average total number of purchased stickers is $7 \cdot 1044.39 \approx 7311$, and the corresponding number of duplicate stickers is about $7311 - 980 = 6331$. Therefore, only about $980/7311 \approx 13\%$ of all purchased stickers contribute directly to album completion, while about 87% are duplicates.

4.5. Comparison Between Exact Recursion, Monte Carlo Simulation, and Mean-Field Approximation

The exact recursive expectation, the Monte Carlo estimate, and the mean-field approximation have different roles in the analysis and should not be interpreted as interchangeable quantities. The exact recursive value $E_0 = 1042.36$ packs is the theoretical expected completion time of the finite-state model under the stated baseline assumptions. The Monte Carlo mean of 1044.39 packs is a numerical estimate obtained from 10,000 independent realizations of the same stochastic process. Its small difference from the exact recursive value is consistent with finite-sample simulation variability.

The mean-field approximation gives 1045.07 packs and is used only as a benchmark. It provides a close numerical approximation for the considered parameter values, but it does not replace the exact recursive formulation because it does not describe the full transition distribution between collection states.

The empirical spread of completion outcomes is interpreted only through the Monte Carlo distribution. Therefore, the favorable and unfavorable completion scenarios reported in this paper are simulation-based percentiles, not exact recursive results and not mean-field estimates.

4.6. Sensitivity to Model Assumptions

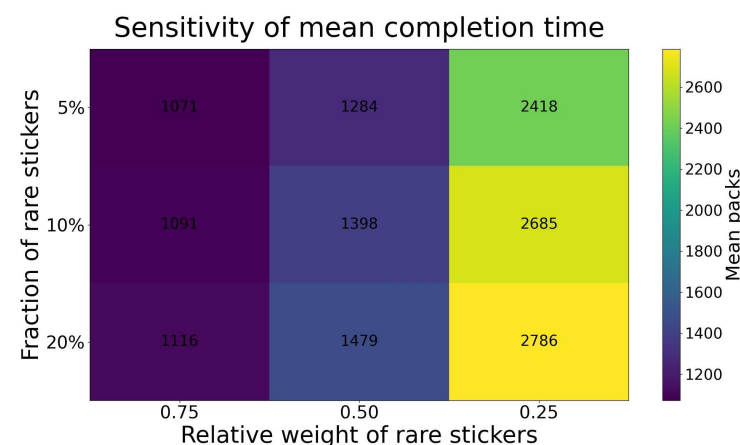
The baseline model is based on three simplifying assumptions: uniform sticker probabilities, independence between packs, and no duplicate stickers within a single pack. These assumptions make the model analytically tractable and define a clear reference scenario. However, real sticker distributions may deviate from this idealized setting. Therefore, the obtained results should be interpreted together with a sensitivity analysis.

The most important possible deviation is non-uniform sticker availability. If some stickers are effectively rarer than others, the final stage of the collection process becomes more difficult, because the expected completion time is then dominated by the least frequently observed stickers. To examine this effect more systematically, additional sensitivity simulations were performed using a two-parameter grid. The first parameter was the fraction of stickers treated as rare, set to 5%, 10%, and 20%. The second parameter was the relative sampling weight of those rare stickers, set to 0.75, 0.50, and 0.25. All remaining stickers retained unit weight. These non-uniform cases are illustrative stress tests and are not intended to represent a measured Panini production distribution. Each sensitivity scenario was evaluated using 10,000 independent simulation runs. The results are summarized in Table 5.

Table 5. Sensitivity of the mean completion time to systematic deviations from the uniform sampling assumption.

Rare Sticker Fraction	Rare Sticker Weight	Mean Packs
5%	0.75	1071.47
5%	0.50	1284.17
5%	0.25	2417.71
10%	0.75	1091.01
10%	0.50	1398.15
10%	0.25	2684.52
20%	0.75	1115.70
20%	0.50	1478.83
20%	0.25	2785.72

The same sensitivity results are visualized in Figure 3. The results show that the expected completion time increases systematically as the rare-sticker fraction increases and as the relative sampling weight of rare stickers decreases. For mild non-uniformity, where 5% of stickers have a relative weight of 0.75, the mean completion time increases only moderately, to about 1071 packs. In contrast, stronger non-uniformity produces a much larger effect. For example, when 10% of stickers have a relative weight of 0.25, the mean completion time increases to about 2685 packs. When 20% of stickers have the same relative weight, the mean completion time is about 2786 packs.

**Figure 3.** Sensitivity of the mean completion time to non-uniform sticker availability. The figure shows the mean number of required packs as a function of the rare-sticker fraction and the relative sampling weight of rare stickers.

This confirms that the completion process is highly sensitive to the lower tail of the sticker-probability distribution. Even a relatively small group of less frequently sampled stickers can dominate the final stage of the acquisition process. Therefore, if real sticker production or distribution were significantly non-uniform, the purely random completion cost could be substantially higher than the value predicted by the uniform baseline model.

This sensitivity analysis does not attempt to infer the true production probabilities of individual stickers because such data are not publicly available. Instead, it shows how the model behaves under controlled deviations from the main assumptions. The baseline result should therefore be interpreted as the expected completion cost under an idealized uniform random acquisition process, while non-uniformity represents a potential source of additional completion cost. For this reason, the uniform random-acquisition result is used as a reference baseline rather than as a mathematical upper bound.

4.7. Practical Implications

From a practical point of view, the obtained results show that the completion of the FIFA World Cup 2026 sticker album requires substantially more packs than one would expect from a simple linear estimate. Although the ideal lower bound is only 140 packs, the expected number exceeds 1000 packs.

This difference is caused by the long-tail stage of the acquisition process. Near completion, most purchased packs contain only stickers that are already present in the album, so the marginal benefit of additional random purchases becomes low. Therefore, random pack purchasing is most efficient in the early and intermediate stages, whereas duplicate exchange and targeted acquisition become increasingly important near completion.

4.8. Cost Analysis and Practical Completion Strategies

For a more practical interpretation of the obtained results, the required number of packs can be translated into an approximate monetary cost in different markets. Based on official retail prices available at the time of writing, a sticker pack costs approximately 1.50 EUR in Europe [16], 1.25 GBP in the United Kingdom [17], and 2.00 USD in the United States [18]. Since retail prices may vary slightly across countries, retailers, and promotional campaigns, the following values should be interpreted as representative market estimates.

Using the exact recursive baseline expectation rounded to approximately 1042 packs for practical cost interpretation, the expected completion cost becomes

- Europe: $1042 \text{ packs} \times 1.50 \text{ EUR} \approx 1563 \text{ EUR}$;
- United Kingdom: $1042 \text{ packs} \times 1.25 \text{ GBP} \approx 1303 \text{ GBP}$;
- United States: $1042 \text{ packs} \times 2.00 \text{ USD} \approx 2084 \text{ USD}$.

These values are based on the exact recursive baseline expectation rounded to a whole number of packs for practical interpretation, whereas the favorable and unfavorable cost scenarios below are based on empirical Monte Carlo percentiles.

Similarly, the favorable scenario of approximately 811 packs corresponds to about 1217 EUR, 1014 GBP, and 1622 USD, respectively, whereas the unfavorable scenario of approximately 1387 packs corresponds to approximately 2081 EUR, 1734 GBP, and 2774 USD.

These values illustrate that completing the album exclusively through random pack purchases may require a substantial financial investment. At the same time, they also explain why purely random completion is rarely the strategy used in practice.

One widely used mechanism for cost reduction is duplicate exchange. Since approximately 87% of all purchased stickers are duplicates, the collector accumulates a large pool of tradable items during the process. By exchanging duplicates with other collectors, the effective number of required purchases can be reduced substantially, particularly near the end of the collection process.

Another common mechanism is the direct purchase of missing stickers. In previous editions, collectors were allowed to order a limited number of specific missing stickers, typically at a unit price higher than the average sticker price obtained through random packs. From the viewpoint of the coupon collector model, such a strategy effectively truncates the most inefficient part of the acquisition process, namely the long-tail phase associated with the final missing stickers.

These realistic collecting behaviors can be represented by a simple hybrid completion strategy. Let r denote the number of stickers that are still missing when the collector stops relying exclusively on random pack purchases. The remaining stickers can then be obtained through duplicate exchange, direct purchase of missing stickers, or secondary-market acquisition. Let $E_0^{(r)}$ denote the expected number of random packs required to reach a state with $n - r$ collected stickers. The value $E_0^{(r)}$ can be computed using the same

finite-state recursion as in Section 2, but with the stopping condition changed from full completion to the state in which r stickers remain missing. The expected cost of such a hybrid strategy can be written as

$$C_{\text{hyb}}(r) = p_{\text{pack}} E_0^{(r)} + r p_{\text{target}}, \quad (15)$$

where p_{pack} is the price of one pack, and p_{target} is the effective average cost of obtaining one remaining missing sticker by a non-random mechanism. This effective cost may represent direct purchase, secondary-market purchase, or the indirect cost of exchange. In the limiting case of successful duplicate exchange without additional payment, the effective value of p_{target} can be close to zero for the exchanged stickers.

The optimal switching point can be determined by minimizing the expected hybrid cost over all possible values of r . Thus,

$$r^* = \arg \min_{0 \leq r \leq n} C_{\text{hyb}}(r). \quad (16)$$

This optimization was evaluated for the European pack price $p_{\text{pack}} = 1.50$ EUR and for several illustrative values of the effective targeted-acquisition cost p_{target} . The values of p_{target} are not intended to represent fixed market prices but, rather, to show how the optimal strategy changes when the effective cost of obtaining a specific missing sticker changes. The obtained optimal switching points and minimum expected costs are summarized in Table 6, where $E_0^{(r^*)}$ denotes the expected number of random packs purchased before switching, and $C_{\text{hyb}}(r^*)$ denotes the corresponding minimum expected hybrid cost.

Table 6. Optimal switching point and minimum expected cost of the hybrid completion strategy for different assumed effective targeted-acquisition costs.

p_{target}	r^*	$E_0^{(r^*)}$	$C_{\text{hyb}}(r^*)$
1 EUR	209	about 216 packs	about 533 EUR
2 EUR	104	about 313 packs	about 677 EUR
4 EUR	52	about 409 packs	about 821 EUR
8 EUR	26	about 504 packs	about 965 EUR

The corresponding hybrid cost curves are shown in Figure 4. The pure random-acquisition baseline cost obtained from the exact recursive expectation is approximately 1564 EUR in the European market. Therefore, all four illustrative hybrid strategies reduce the expected completion cost substantially. The optimal switching point moves closer to album completion as p_{target} increases. In other words, if targeted acquisition is relatively cheap, it is optimal to stop random pack purchases earlier. If targeted acquisition is expensive, random pack purchases remain useful for a longer part of the collection process.

The economic reason for using such hybrid strategies becomes clear when the local marginal cost of random acquisition is considered. If r stickers are still missing, the expected number of newly useful stickers obtained from one additional random pack is

$$E[\Delta m \mid r] = k \frac{r}{n}. \quad (17)$$

Therefore, the approximate random-pack cost per newly useful sticker at that stage is

$$c_{\text{rand}}(r) = \frac{p_{\text{pack}}}{kr/n} = \frac{p_{\text{pack}} n}{kr}. \quad (18)$$

For the European pack price $p_{\text{pack}} = 1.50$ EUR, $n = 980$, and $k = 7$, the corresponding marginal random-acquisition costs are shown in Table 7.

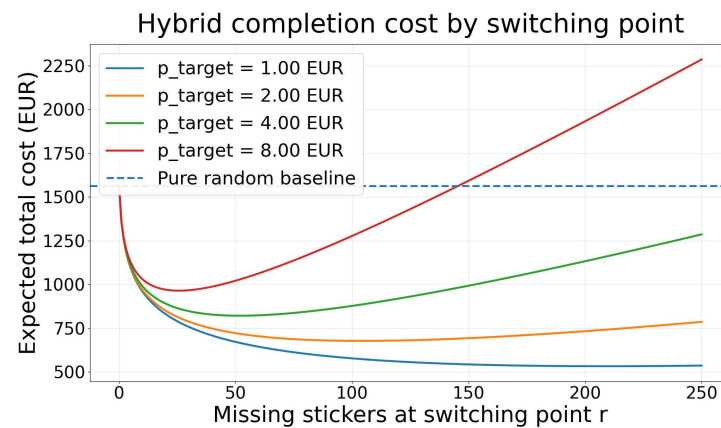


Figure 4. Expected hybrid completion cost as a function of the switching point for different assumed effective targeted-acquisition costs. The minima of the curves correspond to the optimal switching points reported in Table 6.

Table 7. Approximate marginal cost of obtaining one newly useful sticker by additional random pack purchases in the final stage of the collection process.

Missing Stickers r	Expected New Stickers per Pack	Cost per Newly Useful Sticker
100	0.714	2.10 EUR
50	0.357	4.20 EUR
25	0.179	8.40 EUR
10	0.071	21.00 EUR

Table 7 shows why purely random acquisition becomes economically inefficient near completion. Even if a targeted missing sticker is more expensive than the average sticker obtained from a random pack, it can still be cheaper than continuing with random pack purchases in the final stage. Duplicate exchange has a similar effect, because it reduces the number of stickers that have to be obtained through further random purchases.

Therefore, from an optimization point of view, the completion process can be divided into three practical stages. In the initial stage, random acquisition is efficient and should be preferred. In the intermediate stage, duplicate exchange becomes increasingly valuable. In the final stage, targeted acquisition of the remaining stickers becomes the most efficient approach, because it prevents the collector from entering the most expensive part of the purely random process.

In practical terms, this suggests that random pack purchasing is most efficient during the early and intermediate stages of the collection process. As the completion ratio approaches unity, the marginal benefit of purchasing additional random packs decreases rapidly, making alternative strategies considerably more cost-effective.

Thus, the theoretical and simulation-based results presented in this paper should be interpreted as a conditional baseline cost estimate under uniform random acquisition. They should not be interpreted as a strict upper bound on the completion cost. If sticker availability is non-uniform, the completion time and cost may be higher than in the uniform baseline case. Conversely, in real collecting practice, the final cost can be reduced substantially through duplicate exchange, targeted acquisition strategies, and secondary-market transactions.

5. Engineering Case Study: LoED LoRaWAN Dataset

To address the engineering relevance of the considered acquisition structure, this section analyzes an existing real-world LoRaWAN dataset. The aim is not to validate the sticker-album model as a LoRaWAN performance model but to examine whether

analogous first-coverage and redundancy effects can be observed in a real engineering data-acquisition system.

The considered case study is based on the LoRaWAN at the Edge Dataset (LoED), a real-world dataset containing gateway-level LoRaWAN packet traces collected in an urban deployment [9]. LoRaWAN networks are widely used in low-power IoT applications because they support long-range communication with low energy consumption and low device complexity, which is consistent with the broader role of wireless sensor networks, IoT systems, smart-city testbeds, network-discovery problems, and information-freshness-oriented data acquisition in distributed sensing contexts [19–24]. In multi-gateway LoRaWAN deployments, the same uplink transmission can be received by more than one gateway, so the Network Server has to perform packet de-duplication before forwarding the information to the application layer [8]. This makes LoRaWAN data collection a realistic engineering context in which repeated acquisition and redundancy are practically relevant.

For the purpose of the present analysis, the LoED packet traces were processed as an address-level first-coverage problem. Only valid uplink packets were considered. In particular, packets were retained if they had a valid timestamp, a valid device address, a valid physical payload, a positive CRC status, and an uplink message type. Gateway-level duplicate receptions were then removed by grouping repeated physical payloads occurring within a 200 ms de-duplication window. This de-duplication step follows the practical LoRaWAN interpretation that multiple gateway receptions of the same physical uplink should be treated as one transmitted packet rather than as independent observations.

After this filtering and de-duplication step, the full LoED dataset contained 5,413,946 de-duplicated uplink packets and 8503 observed distinct device addresses. The first-coverage process was then evaluated by scanning the de-duplicated packets in chronological order and counting how many distinct device addresses had been observed after each packet. A newly observed device address was treated as a useful first-coverage event, while any later packet from an already observed address was treated as a repeated observation with respect to first coverage. The obtained summary statistics are shown in Table 8.

Table 8. Summary of the LoED address-level first-coverage analysis.

Quantity	Value
Raw gateway-level rows	11,263,001
Valid uplink rows after filtering	8,055,807
De-duplicated uplink packets	5,413,946
Gateway duplicate rows removed	2,641,861
Observed distinct device addresses	8503
Useful first-coverage fraction	0.157%
Redundant-observation fraction	99.843%
Top 10 device-address packet share	74.29%
Top 100 device-address packet share	96.09%
Packets needed for 95% of observed addresses	5,161,176
Packets needed for 99% of observed addresses	5,375,132

The measured first-coverage curve is shown in Figure 5. The curve exhibits the same qualitative long-tail behavior as the sticker-album completion process. At the beginning, new device addresses are observed relatively quickly. Later, the process becomes increasingly dominated by repeated observations of addresses that have already been seen. In the full dataset, 90% of the observed device addresses are reached only after 4,956,599 de-duplicated uplink packets, while 99% are reached only after 5,375,132 de-duplicated uplink

packets. Therefore, the final part of address-level first coverage is strongly dominated by the least frequently observed addresses.

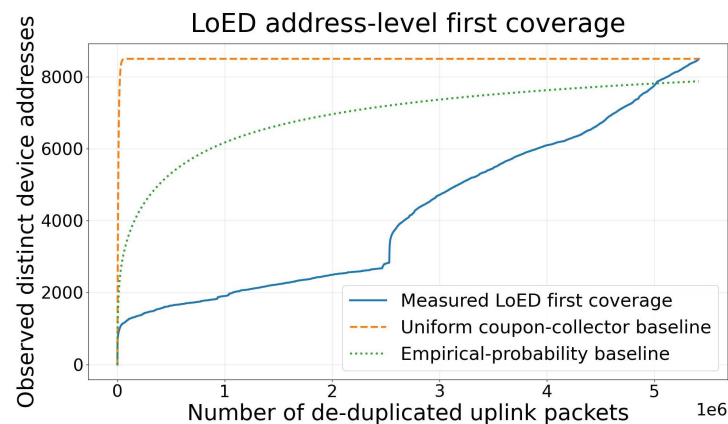


Figure 5. Measured LoED address-level first-coverage curve compared with uniform and empirical-probability coupon-collector baselines. The measured curve shows rapid initial acquisition followed by a long-tail region dominated by repeated observations of already observed device addresses.

For comparison, Figure 5 also shows two coupon-collector-type baselines. The first is a uniform baseline in which all observed device addresses are assumed to be equally likely. The second is an empirical-probability baseline in which the probability of observing a device address is estimated from its relative frequency in the de-duplicated LoED packet trace. If p_i denotes the empirical probability of observing address i , the expected number of distinct observed addresses after N de-duplicated packets is

$$E[C(N)] = \sum_{i=1}^{n_s} \left(1 - (1 - p_i)^N\right). \quad (19)$$

Here, n_s denotes the number of observed distinct device addresses in the analyzed LoED trace. This empirical-probability model is not synthetic because the probabilities p_i are estimated directly from the real de-duplicated packet trace. It is also more appropriate for the LoED case than a uniform model because real IoT traffic is strongly non-uniform. Some devices transmit much more frequently than others, while some addresses appear only rarely.

The non-uniformity of the LoED trace is shown directly in Figure 6, where device addresses are ranked according to their number of de-duplicated uplink packets. The distribution is highly concentrated: the 10 most frequently observed device addresses account for 74.29% of all de-duplicated uplink packets, while the 100 most frequently observed addresses account for 96.09% of them. This explains why the measured first-coverage curve differs strongly from the uniform baseline. The engineering process is not dominated by an average device but by the lower tail of rarely observed addresses.

These results provide a real engineering example of the same qualitative acquisition phenomenon that appears in the sticker-album model. In the sticker case, repeated stickers are duplicates with respect to album completion. In the LoED case, repeated packets from already observed device addresses are redundant with respect to address-level first coverage. In both cases, the acquisition process initially progresses quickly, but the final stage is dominated by rarely acquired entities.

At the same time, the analogy has important limitations. The LoED analysis is based on device addresses, not on independently verified permanent physical device identities. In LoRaWAN systems, a device address is a practical network-level identifier, but it is not always identical to a globally stable physical device identity. In addition, real LoRaWAN

systems include non-uniform device traffic, gateway placement effects, channel-dependent reception, duty-cycle restrictions, retransmissions, adaptive data rate mechanisms, and backend de-duplication rules. Therefore, the sticker-album model should not be interpreted as a validated LoRaWAN performance model.

The value of the case study is instead more specific. It shows, using a real LoRaWAN dataset, that first coverage in engineering data acquisition can exhibit the same qualitative long-tail structure as the sticker-album completion problem. It also supports the sensitivity analysis presented earlier in the paper: non-uniform acquisition probabilities are not merely an artificial modeling assumption but a dominant feature of real IoT packet traces. From an engineering point of view, this motivates practical mechanisms such as packet de-duplication, scheduled reporting, adaptive polling, retransmission control, and the prioritization of devices or addresses that have not yet been observed.

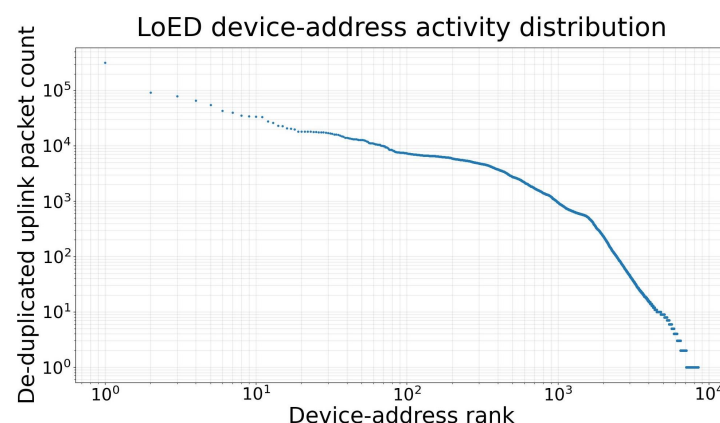


Figure 6. Rank-frequency distribution of de-duplicated uplink packets per observed LoED device address. The strong concentration of packets among a small number of addresses explains the non-uniform first-coverage behavior.

6. Conclusions

In this paper, the completion of the FIFA World Cup 2026 sticker album was analyzed as a multi-item coupon-collector-type acquisition problem. The aim was to estimate, a priori, the expected number of required packs, the total number of purchased stickers, and the corresponding financial cost under a uniform random-acquisition baseline.

The exact finite-state recursive model gives a theoretical baseline expectation of 1042.36 packs under the adopted assumptions. The Monte Carlo simulation gives a finite-sample mean of 1044.39 packs based on 10,000 independent trials and is used to characterize the distribution and variability of completion outcomes. The mean-field approximation gives a similar benchmark value of 1045.07 packs, but it is not treated as the primary theoretical result.

Besides the average value, the obtained results show that the completion process exhibits substantial variability. Based on the empirical Monte Carlo distribution, realistic completion scenarios range from 811 packs in a more favorable case to 1387 packs in a less favorable case. Therefore, the expected value alone is not sufficient for a complete practical interpretation of the problem.

The results also quantify the inefficiency caused by duplicate accumulation. Under the baseline Monte Carlo estimate, approximately 7311 stickers are purchased on average, of which 980 contribute directly to album completion. This corresponds to a useful fraction of about 13%, while the remaining 87% are duplicates. Since the duplicate count follows directly from the number of purchased packs, it was treated as a derived quantity rather than as an independent stochastic output.

When translated into monetary terms, the uniform random-acquisition baseline indicates an expected cost of more than 1500 EUR in a representative European market. However, this value should not be interpreted as a prediction of the actual cost incurred by individual collectors or as a strict upper bound. Non-uniform sticker availability could increase the completion time relative to the uniform baseline, whereas duplicate exchange, the direct purchase of missing stickers, and secondary-market acquisition can reduce the final cost by avoiding the most inefficient long-tail phase of random acquisition.

The sensitivity analysis confirms that the results depend strongly on the assumption of uniform sticker availability. In the considered non-uniform stress tests, the mean completion time increased substantially when a fraction of stickers was assigned a lower relative sampling weight. This shows that the lower tail of the sticker-probability distribution can dominate the completion process if real sticker availability is not uniform.

The hybrid-strategy analysis further shows that purely random acquisition is generally not cost-optimal near completion. For the illustrative targeted-acquisition costs considered in the paper, the optimal strategy is to switch from random packs to targeted acquisition before the album is fully completed. The exact switching point depends on the effective cost of obtaining a specific missing sticker.

The main limitation of the analysis is that empirical data on actual sticker frequencies, real collector exchange behavior, the availability of direct missing-sticker purchase, and secondary-market prices are not available at the time of writing. For this reason, the reported monetary values should be interpreted as conditional baseline estimates under explicitly stated assumptions, not as forecasts of the expenditure of a specific collector. Future work could refine the model by incorporating empirical sticker-frequency data, observed trading behavior, and market-price information once such data become available.

The LoED case study further shows that analogous long-tail acquisition behavior can also be observed in a real engineering data-acquisition system. In the analyzed LoRaWAN packet trace, address-level first coverage is strongly affected by repeated observations and highly non-uniform device-address activity. However, the sticker-album model should not be interpreted as a validated LoRaWAN performance model because the LoED analysis is based on address-level packet traces and not on a complete physical-device model.

Overall, the presented results provide a quantitative baseline for understanding the completion process of a large sticker album under random acquisition. More generally, they illustrate how duplicate accumulation, diminishing returns, and long-tail completion effects can strongly influence the cost and efficiency of practical acquisition processes.

Author Contributions: Conceptualization, P.K.; methodology, P.K. and D.Š.; software, P.K. and D.Š.; validation, P.K. and D.Š.; formal analysis, P.K.; investigation, P.K. and D.Š.; resources, P.K.; data curation, D.Š. and P.K.; writing—original draft preparation, P.K. and D.Š.; writing—review and editing, P.K.; visualization, D.Š.; supervision, P.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the project “Ulaganje u opremanje i adaptaciju istraživačko-inovacijske infrastrukture Sveučilišta VERN” (Investment in research and innovation infrastructure at VERN’ University, NPOO.C3.2.R2-I2.12.0039), funded by the European Union through the National Recovery and Resilience Plan (NextGenerationEU).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The simulation and processing code used in this study is publicly available in the GitHub repository: <https://github.com/DiegoSUM/Coupon-Collector-Problem-Calculator> (accessed on 20 July 2026). The LoRaWAN case study uses the publicly available LoRaWAN at the Edge Dataset (LoED) described in [9]. Due to the size of the raw LoED dataset, it is

not redistributed in the repository; however, the provided processing script can be used to reproduce the derived LoED first-coverage results from the public dataset.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT, GPT-5.5 Thinking, a large language model developed by OpenAI, for the purposes of language refinement and text editing. All scientific content, analysis, and conclusions were developed and verified by the authors. The authors reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. The Cardboard Collection. Complete Guide to Panini World Cup Sticker Albums. Available online: <https://www.cardboardconnection.com/panini-world-cup-sticker-guide> (accessed on 20 July 2026).
2. FIFA. Number of Teams at Every World Cup. Available online: <https://www.fifa.com/en/tournaments/mens/worldcup/articles/number-teams-every-edition> (accessed on 20 July 2026).
3. FIFA. FIFA World Cup 2026: Fixtures, Groups, Teams, Tickets, Host Countries, Cities and More. Available online: <https://www.fifa.com/en/tournaments/mens/worldcup/canadamexicousa2026/articles/fifa-world-cup-2026-hosts-cities-dates-usa-mexico-canada> (accessed on 20 July 2026).
4. Panini: FIFA WORLD CUP 2026™ Official Sticker Collection-Big Collector's Box Panini. Available online: https://paninistore.com/shp_int_en/fifa-world-cup-2026-official-sticker-collection-big-collector-s-box-panini-005460box144oe-it.html (accessed on 20 July 2026).
5. Flajolet, P.; Gardy, D.; Thimonier, L. Birthday paradox, coupon collectors, caching algorithms and self-organizing search. *Discret. Appl. Math.* **1992**, *39*, 207–229. [CrossRef]
6. Sardy, S.; Velenik, Y. Paninimania: Sticker Rarity and Cost-Effective Strategy. Available online: <https://www.unige.ch/~velenik/Vulg/Paninimania.pdf> (accessed on 20 July 2026).
7. Šumberac, D. Numerička Demonstracija i Verifikacija Matematičkog Problema Sakupljača Kupona na Komercijalnom Praktičnom Primjeru. Undergraduate Thesis, VERN' University, Zagreb, Croatia, 2025. Available online: <https://zir.nsk.hr/en/object/vern:3623> (accessed on 20 July 2026).
8. Spadaccino, P.; Crinò, F.G.; Cuomo, F. LoRaWAN Behaviour Analysis Through Dataset Traffic Investigation. *Sensors* **2022**, *22*, 2470. [CrossRef] [PubMed]
9. Bhatia, L.; Breza, M.; Marfievici, R.; McCann, J.A. LoED: The LoRaWAN at the Edge Dataset. In Proceedings of the 3rd International SenSys+BuildSys Workshop on Data Acquisition to Analysis (DATA 2020), Virtual Event, 16 November 2020.
10. Feller, W. *An Introduction to Probability Theory and Its Applications*; Wiley: New York, NY, USA, 1968; Volume 1.
11. Boneh, A.; Hofri, M. The coupon-collector problem revisited—A survey of engineering problems and computational methods. *Commun. Stat. Stoch. Model.* **1997**, *13*, 39–66. [CrossRef]
12. Graham, R.L.; Knuth, D.E.; Patashnik, O. *Concrete Mathematics*, 2nd ed.; Addison-Wesley: Reading, MA, USA, 1994.
13. Fishman, G.S. *Monte Carlo: Concepts, Algorithms, and Applications*, 1st ed.; Springer: New York, NY, USA, 1996.
14. Rubinstein, R.R.; Kroese, D.P. *Simulation and the Monte Carlo Method*, 3rd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2017.
15. Coupon Collector's Problem Calculator. GitHub. Available online: <https://github.com/DiegoSUM/Coupon-Collector-Problem-Calculator> (accessed on 20 July 2026).
16. Offizielle FIFA World Cup 2026™ Stickerkollektion-Box Mit 100 Tüten. Panini Deutschland. Available online: https://www.panini.de/shp_deu_de/offizielle-fifa-world-cup-2026-stickerkollektion-box-mit-100-t-ten-005460box100d-de01.html (accessed on 20 July 2026).
17. FIFA WORLD CUP 2026™ Official Sticker Collection-Box of 100 Packets. Panini UK. Available online: https://www.panini.co.uk/shp_gbr_en/fifa-world-cup-2026-official-sticker-collection-box-of-100-packets-005460box100uk-uk02.html (accessed on 20 July 2026).
18. FIFA World Cup 2026™ Official Sticker Collection-50 Count Box. Panini America. Available online: <https://www.paniniamerica.net/fifa-world-cup-2026-official-sticker-collection-50-count-box.html> (accessed on 20 July 2026).
19. Gubbi, J.; Buyya, R.; Marusic, S.; Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Gener. Comput. Syst.* **2013**, *29*, 1645–1660. [CrossRef]
20. Akyildiz, I.F.; Su, W.; Sankarasubramaniam, Y.; Cayirci, E. A survey on sensor networks. *IEEE Commun. Mag.* **2002**, *40*, 102–114. [CrossRef]
21. Yick, J.; Mukherjee, B.; Ghosal, D. Wireless sensor network survey. *Comput. Netw.* **2008**, *52*, 2292–2330. [CrossRef]
22. Ahmed, Y.; Ghosh, A.; Wang, C.; Shroff, N.B. Beyond Freshness and Semantics: A Coupon-Collector Framework for Effective Status Updates. *arXiv* **2026**, arXiv:2603.26998.

23. Gvozdenovic, S.; Becker, J.K.; Mikulskis, J.; Starobinski, D. IoT-Scan: Network Reconnaissance for the Internet of Things. *IEEE Internet Things J.* **2024**, *11*, 13091–13107. [\[CrossRef\]](#)
24. Sanchez, L.; Muñoz, L.; Galache, J.A.; Sotres, P.; Santana, J.R.; Gutierrez, V.; Ramdhany, R.; Gluhak, A.; Krco, S.; Theodoridis, E. SmartSantander: IoT experimentation over a smart city testbed. *Comput. Netw.* **2014**, *61*, 217–238. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.