HiLL PUBLISHING

Engineering Advances, 2026, 6(4), 238-241
https://www.hillpublisher.com/journals/ea/
ISSN Online: 2768-7961

Enhancing Computation and Data Consistency Optimization in Advertising Click-to-Conversion Chains

Jinxin Li

Electrical & Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA.

How to cite this paper: Jinxin Li. (2026) Enhancing Computation and Data Consistency Optimization in Advertising Click-to-Conversion Chains. *Engineering Advances*, 6(4), 238-241.
DOI: 10.26855/ea.2026.12.001

Received: August 25, 2026
Accepted: September 23, 2026
Published: October 27, 2026

***Corresponding author:** Jinxin Li, Electrical & Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA.

© 2026 by the author(s).
This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives (CC BY-NC-ND) license, which permits non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and is not modified or adapted.
https://creativecommons.org/licenses/by-nc-nd/4.0/

Abstract

Complex multi-terminal and multi-scenario data collection in advertising click-conversion links often leads to disorder of timestamps, mismatch of behavioral sequences, data loss, and duplicate recordings, which result in poor data consistency and inaccurate conversion statistics. Six specific methods have been suggested to overcome these difficulties. By using time series alignment rules based on uniform server clocks, the behavioral time points are adjusted, and disordered data is corrected to ensure the same statistical criteria for the time dimension. A distributed data redundancy and error correction mechanism, which employs multi-node backup storage and intelligent filtering logic, is adopted to avoid data loss and remove invalid abnormal data. The cross-node data synchronisation and verification model standardizes the full-link transmission specifications, achieves real-time data synchronisation and cyclic consistency checks with automatic early warning functions. The dynamic abnormal data recovery method classifies and repairs different types of abnormal data and dynamically adjusts the calibration parameters to eliminate statistical errors and improve the data quality and calculation precision of the advertising conversion links.

Keywords

Ad click-conversion link; data consistency; time series alignment; distributed data optimization

With the increasing variety of digital advertising applications, the click-conversion data is gathered, transmitted, and processed on various terminals, platforms, and server nodes, resulting in a complicated distributed data handling system. The asynchronous clock mechanisms among multi-terminals, the delay in data interactions between different nodes, and the diverse traffic conditions cause many data abnormalities in the advertising behavior records, such as time inconsistency, duplicate entries, data omission, and logical disorder. These abnormal data phenomena affect the uniformity and truthfulness of the entire link data statistics and the correct calculation of attribution, which has a serious influence on the objective assessment of the advertisement effectiveness and restricts the accuracy of the advertisement data analysis. Under such circumstances, it is necessary to establish a comprehensive set of data consistency improvement systems, including time standardization, error correction, synchronization checks, and abnormal recovery. Scientific and unified optimization methods can eliminate the multi-faceted data errors, ensure the stability of the data operation in the distributed computing environment, and offer dependable data assistance for the precise conversion analysis of advertisements.

1. Performance Bottlenecks and Data Problem Analysis of Ad Click-Conversion Link

1.1 Full-Process Data Flow Architecture of the Link

The whole process data flow architecture of the ad click-conversion link includes the complete closed-loop operation from ad display to the final user's conversion and data accumulation. This system depends on the multi-terminal user behavior collection modules, real-time data transmission middleware, cloud computing processing clusters, and offline data warehouses to accomplish the sequential data interaction and processing. At the terminal end, user behaviors such as ad viewing, effective clicking, staying on a webpage, and performing the target conversion generate original log data. Then, the terminal sends the structured and unstructured log data to the real-time computing platform via encrypted transmission protocols for preliminary data cleaning, deduplication, and feature labelling. After processing, the real-time data is divided into two transmission paths, one for real-time adjustment of ad delivery strategies and effect monitoring, and the other for entry into the offline data warehouse for cumulative statistical analysis, user profile matching, and conversion attribution sorting. The whole architecture achieves the integration of front-end behavior capturing, mid-end real-time calculation, and back-end data analysis to support the normal iterative and effect evaluation of the ad conversion business [1].

1.2 Link Computing Performance Bottleneck Analysis

The main obstacles of the ad click-conversion link's computing performance are concentrated in real-time data processing, the capacity to bear peak traffic, and situations requiring multi-dimensional calculation. A large amount of instantaneous click and exposure data generated during the peak period of ad delivery will result in a sudden burden on the real-time computing cluster, and the incomplete scheduling of resources will lead to an increase in the data processing delay and prolong the response time of the link. The multi-dimensional calculation, including user attribute identification, conversion weight allocation, and channel effect ranking, adds to the operational load of the computing system, and the single-threaded processing method of some middleware is not suitable for the high-concurrency data calculation requirements [2]. Furthermore, the data cross-checking link between the real-time computing platform and the offline data warehouse contains redundant calculation steps, which consume extra computing resources and decrease the overall operational efficiency of the link. Moreover, fraudulent clicks and invalid traffic further aggravate this burden, as intelligent recognition models based on deep factorization machines have been developed to identify such abnormal click patterns and reduce the computational load caused by invalid data in the advertising conversion process [3]. An unreasonable resource distribution mechanism in the system results in idle resources in low-traffic periods and a shortage of resources in peak periods, forming persistent performance bottlenecks that affect the real-time performance and accuracy of the ad conversion data.

2. Optimization Methods of Computing Capability for Ad Conversion Link

2.1 Link Layered Parallel Computing Architecture Design

The layered parallel computing structure for ad conversion links performs hierarchical division and parallel reconstruction for the conventional serial data processing method to overcome the low efficiency of single-threaded centralized computation. The structure separates the whole link into an acquisition preprocessing stage, a real-time computing stage, an attribution analysis stage, and an output data stage according to the business processing logic and the computing priority. Each independent stage is deployed with a distributed cluster to accomplish specific computing tasks and can operate simultaneously for multiple data streams without affecting each other [4]. The preprocessing stage deals with the cleaning of logs and data standardization to provide standardized data for the following computing nodes. The real-time computing stage processes large amounts of click and conversion data by using parallel thread groups, while the attribution analysis stage handles the multidimensional weight calculation and user behavior matching independently. The layer-by-layer decoupling and parallel operation can decrease the computing burden on one node and shorten the entire link processing time, thus improving the stability and throughput of the ad conversion data computation.

2.2 Lightweight Disassembly Strategy for Hot Computing Tasks

The lightweight disassembly method for hot computing tasks is used to reduce the computing overhead and the occupied resources of the high-load and high-frequency superposed computing tasks in the ad conversion links.

Tasks like real-time conversion attribution, multi-channel effect statistics, and user portrait labeling usually contain a large number of attribute parameters and redundant checking procedures, which prolong the computation time. According to the business correlation and the computing characteristics, these integrated hot tasks are divided into several fine-grained and low-coupling subtasks. The important subtasks that influence the conversion attribution results are kept for precise full-parameter computation, while the other less sensitive subtasks with simple threshold verification and parameter filtering processing are adopted. The system arranges the decomposed subtasks in order of priority and allocates them to the appropriate lightweight computing nodes for execution. By doing this, the unnecessary resource consumption caused by the full-parameter computation of the hot tasks is avoided, and the response speed of the main computing business is significantly improved in peak traffic conditions.

2.3 Dynamic Computing Acceleration Mechanism Based on Resource Scheduling

The dynamic computing acceleration mechanism makes use of the real-time monitoring of link traffic variations and resource occupation to achieve the intelligent distribution and dynamic adjustment of computing resources for ad conversion links [5]. It constructs a real-time resource perception module that gathers operational data such as the CPU occupancy rate, memory usage, and task queue delay of each computing cluster in real time. According to the peak and trough features of the ad delivery business, the system establishes different resource allocation rules and automatically enlarges the computing node resources and increases the parallel computing channels when encountering bursty high traffic. At the same time, it recovers the idle redundant resources of the low-load nodes in real time and reallocates them to the busy computing links to avoid the coexistence of resource idle and shortage. The dynamic resource matching can keep the computing link operating efficiently under various traffic loads and effectively solve the performance bottlenecks caused by the fixed resource allocation in the conventional computing methods.

3. Data Consistency Optimization Strategies for Ad Click-Conversion Link

3.1 Construction of Link Data Time Series Alignment Rules

The construction of link data time series alignment rules is aimed at solving the problems of unordered timestamps and inconsistency of behavioral sequences in multi-terminal and multi-environmental data collection of ad conversion links. It constructs a unified and standardized system for time measurement. This rule system uses the server standard clock as the common reference for the entire link and reduces the time differences caused by different clock parameters of mobile terminals, personal computers, and third-party delivery systems. We classify the user's advertising behaviors into several time points, such as exposure, click, trigger, page access, and conversion completion, and define different time errors for each behavior based on the business characteristics. The system can arrange the irregular behavioral log data in accordance with the standardized timestamps and correct those records that do not conform to the normal conversion rules. Moreover, the proposed time series rules can adjust to various traffic conditions and effectively standardize the time aspects for the data statistics and attribution analysis of the whole ad conversion link, thereby offering support for the subsequent consistent data analysis.

3.2 Data Redundancy and Error Correction Mechanism in Distributed Scenarios

The data redundancy and error correction mechanism is adapted to the distributed deployment of advertising conversion computing clusters to deal with the data loss and repeated recording errors caused by the interaction among multiple nodes. It uses a multi-node redundant storage method for important advertising behavior data and keeps backup copies of the key log data in different distributed nodes to prevent the permanent loss of data due to the failure of a single node. An intelligent redundancy judgment logic is set to eliminate invalid repeated data that occurs during data transmission and keep only the unique valid behavioral records by means of feature code matching and behavioral attribute verification. Machine learning-based fraud detection has been shown to improve advertising performance by filtering invalid traffic and reducing fraud-induced data distortions in programmatic advertising environments [6]. The internal error correction algorithm can automatically detect abnormal data such as missing fields, disordered parameters, and extreme deviation values in the distributed datasets and restore the preliminary data through the historical data correlation and behavioral logic deduction [7]. This mechanism achieves a balance between data security and data accuracy in the distributed operation and significantly reduces the interference of abnormal data in the link computation.

3.3 Cross-Node Data Synchronization and Consistency Verification Model

The cross-node data synchronization and consistency verification model ensures real-time communication and standardized examination of data transmission among all the computing and storage nodes in the advertising conversion process. It establishes a comprehensive data synchronization protocol to standardize the transmission frequency and docking standards of the front-end collection nodes, real-time computing nodes, and offline storage nodes. Maintaining consistent data standards across distributed advertising computing environments is essential for reliable CTR prediction and conversion attribution, as heterogeneous data sources and inconsistent evaluation protocols can significantly affect the accuracy of downstream analytics [8]. It can perform incremental data synchronization during the peak traffic period and full data synchronization during the low traffic period to maintain real-time data consistency throughout the whole process. The model includes several check indicators such as the consistency of data quantity, the integrity and reasonableness of the behavior logic for cyclic automatic checking of cross-node data. When discrepancies or synchronization delays are detected between the nodes, the model will issue an automatic warning and stop the abnormal data from being used in the conversion calculation until the cross-node data consistency is recovered, thus guaranteeing the accuracy and uniformity of the overall data statistics of the whole process.

4. Conclusion

The series of data consistency improvement measures for advertising click-conversion links can resolve the common data quality issues in distributed and multi-environment advertising data processing. The comprehensive system, including time series matching, distributed error correction, cross-node synchronisation verification, and dynamic handling of abnormal data, controls the entire procedure of advertising behaviour data from collection to computation and statistics. Each improvement method supplements and helps one another to overcome the differences in the time dimension, transmission errors among nodes, and various abnormal data influences in the link data. The adaptability of the improvement system can also meet the changing advertising traffic and operational conditions of nodes. The constructed data consistency assurance system greatly improves the authenticity, uniformity, and accuracy of advertising conversion link data, providing a solid technical basis for further in-depth advertising data exploration, effect assessment, and intelligent advertising strategy modification.

References

[1] Guan C, Liu W, Cheng CY, et al. Mining digital ad contents: finding the sweet spot of authenticity via personality profiling. J Curr Issues Res Advert. 2026;47(1):72-94.

[2] Inoue K, Yoshida M. The impact of linguistic features on CTR in Instagram ads: a study of supplement and cosmetic products. PLoS One. 2026;21(4):e0338313.

[3] Ma D, Wan F. Research on intelligent recognition of ad click fraud based on deep FM heterogeneous integration model. Int J High Speed Electron Syst. 2025;35(5):S0129156425407260.

[4] Gao S, Yang Y. Toward a benchmark for CTR prediction in online advertising: datasets, evaluation protocols and perspectives. Electron Commer Res. 2025;(prepublish):1-50.

[5] Bai J, Geng X, Deng J, et al. A comprehensive survey on advertising click-through rate prediction algorithm. Knowl Eng Rev. 2025;40:e25.

[6] Korbut R, et al. The impact of machine learning-based fraud detection on ad performance in programmatic advertising: an empirical evidence from marketing professionals. Springer; 2026.

[7] Molitor D, Zubcsek PP, Spann M, et al. Does location familiarity increase response to mobile ads? J Interact Mark. 2025;60(2):180-195.

[8] Ma Z, Ke Z, Tang L. DQ-MCTR: Data Quality-Aware Multimodal Click-Through Rate Prediction for Online Advertising. In: 2026 7th International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE). Shenyang: IEEE; 2026. p. 912-917.