International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || Volume 10 – Issue 09 || September 2025 || PP. 01-06

Development of Ranking Systems Based on Engagement Metrics: Hotness Algorithms and User Interest Models

Andreev Georgii[1]
[1]*Bachelor's Degree, Moscow Polytechnic University, Moscow, Russia*

Abstract: The article examines an approach to designing content ranking algorithms based on aggregated behavioral metrics such as views, likes, comments, reading time, CTR, and retention. The hotness formula is analyzed, incorporating the temporal dynamics of engagement, the thematic classification of content, and the behavioral profile of users. The importance of adaptive parameter tuning for different content categories and audience segments is emphasized. Special attention is given to the system's robustness against manipulation and the role of stochastic elements in increasing ranking reliability. The proposed approach contributes to improved personalization, enhanced user engagement, and fairer distribution of attention across digital media platforms.
Keywords: Ranking algorithms, user engagement, hotness, user activity, content typology.

I. Introduction

Modern media platforms based on user-generated content (UGC) face the challenge of automatically selecting and ranking publications based on audience engagement. As the volume of content increases rapidly, traditional moderation and manual curation methods become ineffective, leading to information overload and a decline in user experience. In this context, ranking algorithms based on dynamic behavioral metrics have become essential for evaluating the relevance and value of content in real time.

Hotness-based algorithms go beyond absolute popularity indicators (such as view counts) by incorporating time decay, content typology, and user interaction patterns. This helped rebalance the ranking logic, reducing the advantage older content had simply due to longer exposure. Moreover, aggregating metrics such as reading depth, likes, comments, click-through rate (CTR), and retention enables for the construction of a more accurate model of user interest, with improved robustness against manipulation.

The goal of this article is to analyze the design principles of engagement-based ranking algorithms, to present a generalized hotness formula, and to examine its behavior under different conditions. The paper also explores methods of adapting ranking models to various content categories and user segments, along with mechanisms designed to resist metric manipulation and artificial inflation of visibility.

II. Behavioral Engagement Metrics as the Foundation of Algorithmic Content Ranking

Behavioral metrics reflecting the degree of user engagement with content form the foundation of algorithmic ranking on modern digital platforms. These metrics enable a quantitative assessment of audience interest in publications, indirectly capturing their perceived relevance and informational value. To construct a well-founded hotness model, metrics are selected based on several criteria: measurement stability, sensitivity to changes in user behavior, interpretability, and resistance to manipulation (table 1).

Table 1: Characteristics of user engagement metrics [1, 2]

Metric	Type of engagement	Impact on ranking	Advantages	Limitations
Views	Surface-levelinterest.	High (baselinecomponent)	Easy to track; measures reach.	Susceptible to artificial inflation; lacks depth.
Likes	Emotionalapproval.	Medium	Simple signal; user-friendly.	May begivenoutofcourtesyorautomatically.
Comments	Cognitive involvement.	High (withdelay)	Reflectsreflection anddiscussion.	Topic-dependent; influencedbypolarity.
Readingtime	Depthofcontentconsumption.	High	Indicatesrealinterest.	Requiresprecisetrackingmechanisms
CTR	Attractivenessoftitle/preview.	Medium	Capturesbehavioralchoice.	Can be artificially inflated by clickbait
Retention	Sustainedattention	Medium–high	Reflects return	Difficult to apply to one-time

	overtime.		visits and content quality.	publications

The comparative analysis presented in the table shows that each metric captures a different level of user interaction – from superficial attention to deep cognitive engagement. It is important to note that no single metric serves as a universal indicator of interest; their relevance varies depending on the content context, audience profile, and distribution channel. Therefore, building a robust ranking algorithm requires an aggregated model in which metric weights are determined empirically and adapted to the platform's specific characteristics. This approach forms the foundation of the hotness concept, which will be examined in the following section.

III. Content Relevance Modeling: Formalization of the Hotness Metric and Its Temporal Dynamics

The algorithmic evaluation of content relevance and appeal in digital media is increasingly based on an aggregated metric known as hotness. This metric represents a weighted function of multiple behavioral parameters, capturing both the intensity of user interaction and the temporal dimension – that is, the dynamics of engagement over time. The hotness model enables the system to differentiate content not only by its overall popularity, but also by the rate of interest growth, the persistence of attention, and its contextual relevance at a given moment [3].

The general form of the hotness formula can be expressed as follows:

$$H(t) = \sum_{i=1}^{n} w_i \cdot M_i(t) \cdot e^{-\lambda_i (t - t_0)} \qquad (1),$$

Where:

$H(t)$ – is the value of the hotness metric at time t,

$M_i(t)$ – is the value of the i-th engagement metric (e.g., views, likes, comments),

w_i – is the weight coefficient representing the relative importance of the given metric,

λ_i – is the exponential decay coefficient,

t_0 – is the publication timestamp of the content.

For example:

```
SELECT id, title,
  (0.3 * views + 0.5 * likes + 0.2 * comments) *

EXP(-0.0003 * LEAST(extract(epoch from (now() - date_published)), 1000000)) AS hotness -- limiting attenuation 1000000 seconds
FROM posts ORDER BY hotness DESC limit 100;
```

The inclusion of the exponential decay function $e^{-\lambda_i (t - t_0)}$ makes it possible to limit the influence of «older» user interactions and to emphasize recent activity – an especially important consideration in news-driven or fast-changing content streams. As a result, publications that rapidly accumulate engagement within a short time frame are prioritized over more stable but temporally outdated materials.

The context of the platform is used for selecting the measure of engagement M_i. For instance, in a recommendation for news, likes and CTR are assigned higher priority, while comments and reading time are assigned priority in a social network. The weight coefficients w_i can be determined empirically based on historical data on engagement and click-through performance. Moreover, the decay parameter is often determined empirically based on historical data regarding user interaction patterns and content performance.

Moreover, the decay parameters λ_i may vary across metrics: views may decay faster than comments, while likes may persist longer than CTR. This flexibility enables fine-grained control over the «lifespan» of different engagement signals within the model.

Adjusting model parameters can substantially alter the behavior of the ranking algorithm. For example, increasing λ_i accelerates the «aging» of content, favoring newly published materials. Conversely, lowering this coefficient prolongs content visibility by increasing temporal inertia. Similarly, increasing the weights w_i for specific metrics can shift the algorithm's focus toward forms of engagement that are less susceptible to artificial inflation – such as reading depth or user retention.

Example: suppose a publication receives 5,000 views, 300 likes, and 80 comments in the first two hours. With weights $w_{views} = 0.3$, $W_{likes} = 0.4$, $W_{comments} = 0.6$ nd with a decay coefficient $\lambda = 0.5$ the hotness score six hours after publication may be lower than that of another post with only 2,000 views but a sustained discussion and high CTR, depending on how weights and decay parameters are calibrated.

Thus, the hotness formula serves as a core instrument for building fair, adaptive, and robust ranking algorithms under conditions of high-frequency content generation. It provides a quantitative framework not only for assessing content demand but also for evaluating its temporal relevance – specifically, the capacity of content to maintain audience engagement over time.

IV. User Segmentation and Adaptation of the Ranking Algorithm to Audience Types

The effectiveness of a ranking algorithm is significantly influenced when it takes into account the diversity of user behavior patterns and profiles. Users of digital media platforms are heterogeneous, with differences being exhibited in the frequency of engagement with the resource, depth of content usage, participation level, and the motivational dimension. Ignoring such differences translates into loss of relevance in providing personalized content and loss of interest by the audience.

For example, as a 2024 PewResearch survey illustrates, the overwhelming majority of American adults (86%) say they receive news occasionally on a smartphone, computer, or tablet, and 57% say they do this often (fig. 1).

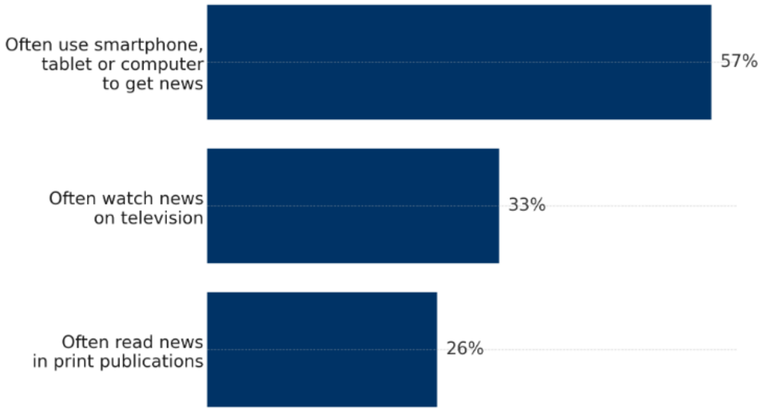

Figure 1: Frequency of news consumption by device type among internet users [4]

These figures emphasize that the vast majority of users would prefer receiving news through electronic devices, hence making the web environment a primary channel for transferring advertising messages. In contrast to television and print media, web and mobile environments provide more immersive and personal contact with the audience, offering extensive possibilities for contextual and native advertising, especially in the form of UGC and natively placed integrations within consumed content.

In the Russian segment, according to Mediascope, there is a growing audience for social media, video platforms, and messengers. The average time spent on the internet was 4 hours and 23 minutes in 2023, 4 hours and 30 minutes in 2024, and 4 hours and 33 minutes in the first quarter of 2025 (fig. 2).

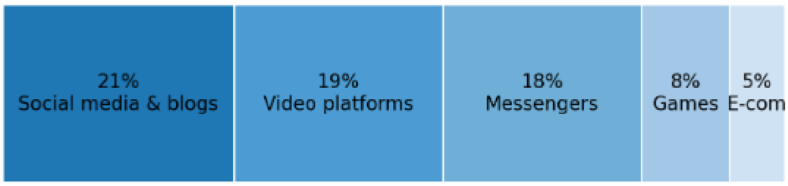

Figure 2: Share of time users spend on different types of online platforms [5]

The growth of digital consumption confirms the increasing role of algorithmic systems in managing user attention. In particular, the rise in average time spent online indicates a saturated media environment and the increasing complexity of behavioral trajectories. This makes the analysis of user audience structure especially relevant and necessitates the implementation of adaptive ranking mechanisms.

One practical approach to modeling user activity is segmentation based on the frequency of returning to the platform. Within the proposed model, at least four stable categories can be identified:

- Daily active users – the core audience of the platform, characterized by consistent engagement and high sensitivity to the quality of content delivery;
- Periodically returning users (once every few days or weekly) – focused on curated «best of the period» content and responsive to delayed engagement signals;
- Newcomers – new users whose behavioral model has not yet been established and who require actively engaging content presentation;
- Dormant users or those returning after a break of more than 14 days – often require reactivation of interest through recommendation mechanisms and exposure to socially significant content.

Depending on the user's segment, the system can adapt the ranking algorithm's parameters. The specific logic of content delivery for different behavioral scenarios is presented in table 2.

Table 2: Adaptation of ranking logic based on user segment

User segment	Priorityincontentdelivery	Algorithmbehavior / parameteradaptation
Daily activeusers	Current, fresh, and widely discussed publications.	Increase in temporal decay coefficient; higher weights for immediate engagement metrics (comments, likes).
Periodicallyreturningusers	Selection of top publications over a recent period.	Decrease in decay rate; emphasis on aggregated engagement across an extended time window.
Newcomers	Publications with high visual and meta-attractiveness (CTR, likes).	Reduction of time decay impact; increased weight for universally engaging metrics.
Dormant / returningusers (>14 days)	Content with high social or news relevance.	Use of modified scoring with emphasis on external signals (author reputation, social relevance).

Such segmentation not only allows for more personalization of content but also resolves the interests of different audience segments. Frequency and behavior attributes-based segmentation allow tuning of ranking mechanisms more precisely, which subsequently boosts the relevance of content displayed, increases user interaction, and optimizes the overall stability of the content system.

V. Content Classification and the Influence of Thematic Affiliation on Algorithmic Visibility

Ranking algorithms in systems with both user-generated and editorial content inevitably face the challenge of content heterogeneity. Different types of publications generate distinct engagement patterns, and a universal model that does not account for thematic specificity may exhibit bias or unjustifiably suppress the visibility of certain content categories [6]. Therefore, an important direction in algorithm adaptation is its configuration based on content typology – a structured differentiation of publications according to topic and purpose (table 3).

Table 3: Content typology and parameter adaptation in ranking algorithms [7, 8]

Content category	Behavioralcharacteristics	Prioritymetrics	Hotnesscalculationfeatures	Limitations / risks
News	Peak activity within the first hours; high short-term engagement	Views, CTR	High decay coefficient, emphasis on recency.	Rapid obsolescence; low sensitivity to comments.
Personal stories	Emotional engagement; sustained discussions and long reading	Comments, retention	Increased weight for long-term metrics; slower decay.	Topic-dependent; subjectiverelevance.

	sessions			
Commercial / nativeads	High CTR; brief user attention; potential for optimization or manipulation.	Readingtime, views, CTR	Down weighted CTR; additional filtering of anomalous activity.	Prone to clickbait; low depth of engagement.
Sociallyoriented	External traffic; limited on-platform measurability.	Outboundclicks, partiallylikes	Indirect evaluation; reduced weight of internal metrics	Attribution difficulties; limited system oversight.

From the algorithm's perspective, each of these content types requires a different weight profile in the calculation of the hotness metric. For example, in the case of news content, it is important to account for exponential decay, whereas for personal stories, comments and reading time become more relevant indicators. In turn, commercial content may be prone to manipulation through the optimization of visual elements, which necessitates lowering the weight of CTR or applying additional filtering of behavioral anomalies.

Adjustment of the hotness formula based on content type can be implemented through parametric tuning of weights and decay coefficients, or by selecting from a set of pre-trained models specific to each content category. For instance:

$$H_c(t) = \sum_{i=1}^{n} w_i^{(c)} \cdot M_i(t) \cdot e^{-\lambda_i^{(c)}(t - t_0)} \tag{2}$$

(2),

Where the index *c* indicates the content category, and the parameters $w_i^{(c)}$ and $\lambda_i^{(c)}$ reflect its specific characteristics.

Such a modification makes it possible to account for differences in audience behavioral responses and enhances the fairness of visibility distribution across content formats. As a result, content representation on the main page becomes more balanced, distortions caused by presentation-specific factors are minimized, and conditions are created for the organic growth of less «noisy» but meaningful publications.

VI. Algorithmic Robustness against Manipulation: Behavioral Anomalies and the Role of Stochastic Elements in Ranking Systems

Ranking systems based on behavioral engagement metrics are inherently vulnerable to intentional data manipulation by users seeking to artificially increase the visibility of their content. This issue is particularly pronounced in commercially motivated publication instances, where authors consciously attempt to reverse-engineer the logic of the algorithm and modify either their content or user behavior to align with the algorithm's rank signals. Certain of the more prevalent manipulative strategies include artificially inflating view metrics (e.g., by bots or coordinated visits), generating shallow engagement (e.g., low-effort comments), and increasing CTR through incendiary or misleading titles (clickbait).

Such artificially induced activity can potentially drive a short-term inflation of the numerically calculated hotness score in the absence of genuine audience interest. Unless managed, these activities undermine the purity of the ranking mechanism and the platform's content discovery pipeline trust. Maintaining stability of the ranking algorithm against such manipulation is therefore the matter of utmost priority.

One effective strategy to enhance algorithmic resilience is the introduction of stochastic variability – for example, randomizing the subset of metrics used in hotness calculation or dynamically adjusting metric weights and decay coefficients within a defined range. This approach increases the unpredictability of the ranking model and makes reverse-engineering more difficult for actors attempting to game the system. In practice, this can involve alternating between scoring user segments or eliminating input from suspected sessions, based on temporal and behavioral profiles (e.g., unusual activity during off-peak times or repeated sequence of interactions).

Another layer of protection involves detecting behavioral anomalies. This is achieved by identifying deviations from normative interaction patterns, such as sharp increases in views without corresponding increases in reading time, an excess of likes with minimal click activity, or templated and low-effort comments. They can trigger remedial reactions – either by down-weighting specific measures or by using an alternative implementation of the formula for hotness.

In summary, algorithmic resilience is maintained through a blend of behavioral anomaly detection, adaptive parameter adjustment, and stochastic modeling. Collectively, these methods guarantee the purity of the ranking system while guaranteeing openness and scalability of the platform.

VII. Conclusion

With the backdrop of the runaway growth of user-generated and editorial content on digital platforms, strong ranking mechanisms have become an essential tool for maintaining the integrity of user experience. The hotness-based ranking model outlined in this article shows how the combination of behavioral signals – temporal dynamics, content form factor, and user segmentation – enables the construction of an adaptive and sustainable content feed. The use of random factors and defensive functions against manipulation additionally strengthens the algorithm's validity, diminishing the influence of fraudulent promotion methods. Thus, the proposed methodology assists in higher relevance of presented content, increased audience activity, and sustainable growth of media websites with a high degree of user involvement.

References

[1] S. Milli, M. Carroll, Y. Wang, S. Pandey, S. Zhao, and A. D. Dragan, Engagement, user satisfaction, and the amplification of divisive content on social media, *PNAS Nexus*, *4(3)*, 2025, pgaf062. DOI: 10.1093/pnasnexus/pgaf062. (8)

[2] O. Nikitenko, The role of CRM systems and digital technologies in optimizing order management and enhancing operational efficiency in the packaging industry, *Slovak International Scientific Journal*, *92*, 2025, 37–39. (8)

[3] G. Song, Y. Wang, X. Chen, H. Hu, and F. Liu, Evaluating User Engagement in Online News: A Deep Learning Approach Based on Attractiveness and Multiple Features, *Systems*, *12(8)*, 2024, 274. DOI: 10.3390/systems12080274. (8)

[4] Pew Research Center, More than eight-in-ten Americans get news from digital devices, 2021. Available at: https://www.pewresearch.org/short-reads/2021/01/12/more-than-eight-in-ten-americans-get-news-from-digital-devices/(accessed 15.08.2025). (8)

[5] Mediascope, Internet audience overview. Available at: https://mediascope.net/library/presentations/ (accessed 15.08.2025). (8)

[6] A. Makhtibekov, Effective approaches to developing digital marketing strategies, *Eurasian Union of Scientists. Series: Technical, Physical and Mathematical Sciences*, *1(3(128))*, 2025, 11–14. (8)

[7] C. Peukert, A. Sen, and J. Claussen, The editor and the algorithm: Recommendation technology in online news, *Management Science*, *70(9)*, 2024, 5816–5831. (8)

[8] R. Ulloa and C. S. Kacperski, Search engine effects on news consumption: Ranking and representativeness outweigh familiarity in news selection, *New Media & Society*, *26(11)*, 2024, 6552–6578. (8)