

Ranking of Social Media Alerts with Workload Bounds in Emergency Operation Centers

Hemant Purohit
George Mason University
Fairfax, VA, USA
hpurohit@gmu.edu

Carlos Castillo
Universitat Pompeu Fabra
Barcelona, Spain
chato@acm.org

Muhammad Imran
Qatar Computing Research Institute
Doha, Qatar
mimran@hbku.edu.qa

Rahul Pandey
George Mason University
Fairfax, VA, USA
rpandey4@gmu.edu

Abstract—Extensive research on social media usage during emergencies has shown its value to provide life-saving information, if a mechanism is in place to filter and prioritize messages. Existing ranking systems can provide a baseline for selecting which updates or alerts to push to emergency responders. However, prior research has not investigated in depth how many and how often should these updates be generated, considering a given bound on the workload for a user due to the limited budget of attention in this stressful work environment.

This paper presents a novel problem and a model to quantify the relationship between the performance metrics of ranking systems (e.g., recall, NDCG) and the bounds on the user workload. We then synthesize an alert-based ranking system that enforces these bounds to avoid overwhelming end-users. We propose a Pareto optimal algorithm for ranking selection that adaptively determines the preference of top- k ranking and user workload over time. We demonstrate the applicability of this approach for Emergency Operation Centers (EOCs) by performing an evaluation based on real world data from six crisis events. We analyze the trade-off between recall and workload recommendation across periodic and realtime settings. Our experiments demonstrate that the proposed ranking selection approach can improve the efficiency of monitoring social media requests while optimizing the need for user attention.

Keywords—Human-centered Computing, Information Overload, Attention Budget, Disaster Management, Pareto Optimality

I. INTRODUCTION

Social media analytics has become a mainstream part of organizational workflows and services in all kinds of organizations, including governments and for-profits. The use of social media in organizations has demonstrated improvements in their customer relations and services. Likewise, for emergency management, a substantive body of research has shown how response agencies and nonprofits can monitor social media for situational awareness [1], [2].

However, due to the characteristics of “big crisis data” [3], which includes high volume and velocity, there are many challenges in monitoring social media message streams. These messages have varied degrees of information and noise that may not be of potential value for the response, ranging from actionable requests or offers of help [4], [5] and unsubstantiated rumors [6] to gratitude and advertisement [7], [8]. Thus, finding the relevant social media updates is a critical concern for emergency management services.

Existing alert-based systems [9], [10] can provide a solution to generate an alert every time a relevant (sub-) event

TABLE I
MOTIVATION TO STUDY A RELATIONSHIP BETWEEN RANKING METRICS AND WORKLOAD. THE DESIGN EXPECTATION OF HIGH RECALL OF RANKED ALERTS WITH A LOW WORKLOAD OF MONITORING IMPLIES THE LESSER NEED OF ATTENTION FROM THE TIME-CRUNCHED END USERS.

		<u>Workload</u>	
		<i>High</i>	<i>Low</i>
<u>Recall</u>	<i>High</i>	inefficient	desired
	<i>Low</i>	worst	ineffective

is detected through social media. However, given a highly dynamic situation, too many alerts can be triggered during a short time window. Thus, such alerts intended to help in social media monitoring would rather serve as distractions to the multitasking EOC personnel and hamper their work during a time-critical event [11], [12]. Information Retrieval literature provides different types of ranking systems [13], which could be employed to select the multiple relevant items for various emergency services. For instance, periodic top- k retrieval systems can be used for batch-based processing of social media, while continuous systems with push-based and pull-based processing can be used for summarizing trends and sending alerts [10], [9], [14]. In particular, push-based alert systems are relevant in the context of an emergency operation center (EOC). However, it is unclear if a “one-size-fits-all” solution of any alert system would be applicable, given the highly stressful work environment of emergency workers that leads to a limited budget of user attention/workload. Table I shows the trade-off and the design expectation for the relationship between the system and user workload metrics. An adaptive alert system with high recall and low required workload would be efficient and effective to minimize the waste of user efforts.

Contributions. We formulate a novel problem of how to create an alert ranking system that is adaptive to the bounds on user performance, for deciding how many alerts to generate and when. We present a novel model of the human-machine interaction to quantify the relationship between the performance metrics of an alert ranking system (e.g., recall) and a user (e.g., workload). We also evaluate, using real-world crisis datasets, an alert ranking system that enforces a maximum user workload to avoid overwhelming users and adaptively select an appropriate set of ranked messages.

The rest of this paper first describes related work in Section II. Section III presents our approach to model the ranking-workload relationship, and an optimization method for selecting a ranking policy. Finally, in Section IV we demonstrate the applicability of this model by analyzing datasets from 6 crises, before discussing the limitations in Section V and our conclusions in Section VI.

II. RELATED WORK

Literature on social media in emergency management is vast, for a survey, see [8], [3]. In this section we focus on works that are closely related to our problem of selecting which ranking system for emergency services is appropriate for operational efficiency.

A. Social media during emergencies

Improving social media-based emergency response is challenging due to “big data” characteristics of high volume, variety, and velocity of social media streams, which overwhelm response services in processing data for relevant information [3]. Crisis informatics [15] research has investigated the use of social media during disasters for response services. The prior research identifies a key challenge and a barrier for the efficient use of social media communication channel for response organizations as the information overload on responders [11], [3]. Information overload originates from a variety of factors including the large scale of unstructured and noisy nature of social media streams and the lack of time to cognitively process relevant content in social media to prioritize for response. In the emergency management domain, Public Information Officers (PIOs) play the crucial role to collect relevant information from public sources for the response agencies or an EOC, by leveraging various information communication technologies including social media [2]. In this context, PIOs of formal emergency management services have started using social media channels to communicate effectively with public and source relevant information for actionable intelligence. The recent reports and surveys of formal response organizations [16], [17] recognize social media as a novel information channel for improving operational response coordination. However, a research question of how and when to effectively monitor social media for finding relevant information remains open.

B. Push and Pull Ranking Systems

We briefly summarize various ranking methods for the problem of alert generation in different domains.

Researchers have investigated diverse techniques for push-based alert systems within the context of disaster management, in particular, using event detection approaches. Sakaki et al. [18] proposed a method for near-realtime earthquake event detection and alert dissemination via Twitter. They devised a classification algorithm to monitor tweets for detecting a target event and send out corresponding alert. Earle et al. [19] also evaluated an earthquake detection and alert dissemination procedure solely relying on temporal pattern

analysis for the keyword-based tweet-frequency time series. Avvenuti et al. [9] developed a burst detection algorithm to promptly identify outbreaking seismic events and automatically broadcasted alerts via a dedicated Twitter account and by email notification systems. Robinson et al. [20] and Yin et al. [21] developed Emergency Situation Awareness platform for earthquake detection in Australia and New Zealand regions using Twitter, which sends email notifications for evidence of earthquakes. Researchers have also designed news feed systems that give top- k alerts, which are relevant for users who subscribe to specific information sources. For instance, Bao and Mokbel [22] developed a location-aware system for news feed ranking, where top- k news feeds were selected based on spatial-temporal proximity and the user preference characteristics. The key limitation of the existing alert generation methods within our problem context is that it is not clear when to generate alerts and how many to generate, for efficiently assisting and not obstructing an EOC expert’s task.

C. Summarization Update Systems

Another major category of work related to our problem is in the area of data stream summarization. Several researchers have devised methods to generate summarization updates for dynamic top- k relevant items. Aslam et al. [10] defined an information access problem in the context of streaming data and proposed a track in the well-known TREC Challenge. The challenge was to develop systems for efficiently monitoring the information associated with an event and broadcast short, relevant, and reliable sentence-length updates about the developing event. Kedzie et al. [23] presented a system for update summarization that predicts the salience of sentences with respect to an event using disaster-specific features including geo-locations and language models, and then bias a clustering algorithm for sentence selection for updates. McCreadie et al. [24] developed a novel incremental update summarization approach that adaptively alters the volume of content issued as updates over time with respect to the prevalence and novelty of discussions about the event. Rudra et al. [25] proposed a framework that first classifies tweets to extract situational information and then, summarizes the information for a user. Their approach factored in the disaster-specific tweet characteristics that contain both situational and non-situational information. Nenkova et al. [26] provide an extensive survey on automated summarization methods. Our goal is to not develop update-summarization algorithm, but the selection policy for adapting the appropriate behavior of the ranking algorithm for updates as per the end user’s workload.

Gap summary: While the above-discussed works on alert generation and stream/update summarization methods are related, they do not account for and study the relationship between the number of (top) ‘ k ’ alerts/updates to generate and the bounds on the user’s workload. Thus, none of the existing systems can be directly used to address our problem. Instead, these works motivate the novel problem to design a generalizable ranking selection method that is aware of the user workload bounds.

TABLE II
EXAMPLE OF A RW MATRIX (ROWS AS k , COLUMNS AS t_{ij} , AND CELLS
AS THE ATTAINABLE RECALL AND WORKLOAD TUPLES).

k	$t_{ij}=10$	$t_{ij}=20$	$t_{ij}=30$	$t_{ij}=40$	$t_{ij}=50$	$t_{ij}=60$
1	(48, 6)	(34,3)	(27,2)	(23,1.5)	(20,1.2)	(18,1)
2	(70,12)	(53,6)	(44,4)	(38,3)	(34,2.4)	(30,2)
...
10	(99,60)	(96,30)	(90,20)	(84,15)	(78,12)	(73,10)

III. APPROACH: WORKLOAD-BOUNDED ALERT RANKING

This section first describes our problem formally and then, the solution to model the ranking-workload relationship as well as select a ranking policy for generating alerts.

Problem Statement. Let t_{ij} be a finite time period from timestamp t_i to t_j ($0 \leq i < j \leq n$), x_{ij} be a finite set of messages generated in t_{ij} , $w(k, t_{ij})$ be the required user workload to monitor k messages ($0 < k < |x_{ij}|$) in t_{ij} , and B as the bound on maximum user workload in the total time period $[t_0, t_n]$, i.e. $\sum_{ij} w(k, t_{ij}) \leq B$. Select a ranking function $R(x_{ij})$ to retrieve top- k items in t_{ij} for alerts such that the ranking-performance metric $M(R(x_{ij}))$ is maximum and the required user workload $w(k, t_{ij})$ is minimum.

Solution. Given the varied types of tasks in EOCs, an alert could be generated for serving different user roles. For a concrete demonstration of our proposed solution to the above problem, we consider the alerts targeted for public communication personnel when a citizen requests to help for a resource or seek information during a disaster. Our solution approach involves three specific steps as described next: i.) relevant message identification and ranking, ii.) ranking-workload (RW) matrix generation, and iii.) optimal RW policy selection.

A. Relevant Message Identification & Ranking

We consider a general class of emergency service requests as relevant messages for alerts that include *actions*, such as a request for resources (e.g., emergency medical assistance for an injured person) as well as *information* (e.g., a request for a phone number for information on missing people) [27], [7]. We have considered serviceability of messages as the relevance criterion [7]. The key characteristic of serviceability of a request message for an alert is that it requests a resource that can be provided, or asks a question that can be answered by the service personnel. Our approach requires a relevancy classification and ranking for the messages. Thus, we adapted the learning-to-rank methodology [28] and designed a SVM-Rank classifier, using the labeled messages with binary relevance classes provided by the emergency domain experts in the prior research [7]. For features, we first used Bag-of-Words features that achieved accuracy of only 65%. Therefore, we resolved to an improved approach for the relevancy classification with better accuracy from our prior work [7] that used additional features of informative details, such as time, place, or context in the message content. Using the relevancy classification and ranking prediction for messages, we compute the ranking metrics for a given set of messages x_{ij} in a period t_{ij} for different types of ranking from top-1 to top- k alerts.

B. Ranking-Workload (RW) Matrix Generation

We propose a matrix-based model to formalize the relationship between ranking performance metrics and the end user workload. We define a RW matrix, as shown in table II, where rows represent the number of top- k alerts to generate and the columns represent the period t_{ij} for the frequency of generating the top- k alerts. The matrix contains 2-tuple values of functions corresponding to the ranking metric and user workload as follows:

$$RW(k, t_{ij}) = \langle M(R(x_{ij})), w(t_{ij}, k) \rangle \quad (1)$$

where,

- $M(R(x_{ij}))$ is the ranking metric function that computes the performance score for a chosen top- k alert ranking $R(x_{ij})$ of message set x_{ij} in t_{ij} , such as Precision@ k , NDCG@ k , and Recall@ k [13].
- $w(t_{ij}, k)$ is the user workload function that characterizes the notion of the amount of hourly work in industry. We define the user workload as the number of alerts to monitor in h hours (or $h * 60$ minutes):

$$w(t_{ij}, k) = k * (h/t_{ij}) \mid 1 \leq t_{ij} \leq h \quad (2)$$

For simplicity, we consider $h = 1$ hour, i.e., 60 minutes and $t_{ij} \in \{10, 20, 30, 40, 50, 60\}$ minutes. For instance, $t_{ij} = 10$ and $k = 5$ imply that an end user will need to monitor top-5 alerts every 10 minutes and the required workload for him will be the cognitive processing of 30 messages per hour.

We consider the top- k alert systems for $k \in [1, 2, \dots, 10]$. We constructed the RW matrix for t_{ij} using the ranking metric function as Recall@ k for the top- k results from the predicted relevant messages in x_{ij} .

C. Optimal RW Policy Selection for Recommendation

Given the multiple choices of workload and desired recall values in the RW matrix, as illustrated in table II, it is challenging to determine which combination of the top- k ranking and t_{ij} period be recommended. For instance, (row $k=1$, column $t_{ij}=60$) shows the minimum workload setting (18,1), although with low recall, while (row $k=10$, column $t_{ij}=10$) shows the maximum recall setting (99,60) with high workload. Thus, maximizing recall for selecting the ranking of top- k alerts may not always lead to the low workload recommendation. It is a multi-objective optimization challenge.

We design our optimization solution using *Pareto Optimality* principle [29], given the lack of ground-truth data and knowledge during the time-critical times about the domain user preferences, which are often required to reach the best solution for multi-objective problems. Our two competing objectives are to achieve low workload and high recall (or low error rate) as illustrated above. An optimal solution would be Pareto optimal when it is not feasible to improve an objective without a penalty to another – a *non-dominating* solution. Formally, a vector of feasible decision variables \hat{x}^* is Pareto optimal if there does not exist another feasible decision vector \hat{x} such that $f(\hat{x}) \leq f(\hat{x}^*)$ and $f'(\hat{x}) < f'(\hat{x}^*)$ for at least one f' .

TABLE III

SUMMARY OF DATASETS FOR TWEETS CONTAINED IN THE DIRECTED CONVERSATIONS TO RESPONSE AGENCIES ON SOCIAL MEDIA, WHICH MAY CONTAIN A POTENTIAL ALERT.

Event (start-end month/day)	Tweets	Relevant	Non-Relevant
Hurricane Sandy 2012 (10/27-11/07)	1,153	40%	60%
Oklahoma Tornado 2013 (05/20-05/29)	1,513	48%	52%
Alberta Floods 2013 (06/16-06/16)	2,727	28%	72%
Nepal Earthquake 2015 (04/15-05/15)	2,222	18%	82%
Louisiana Floods 2016 (10/11-10/31)	1,369	34%	66%
Hurricane Harvey 2017 (08/29-09/15)	12,742	20%	80%

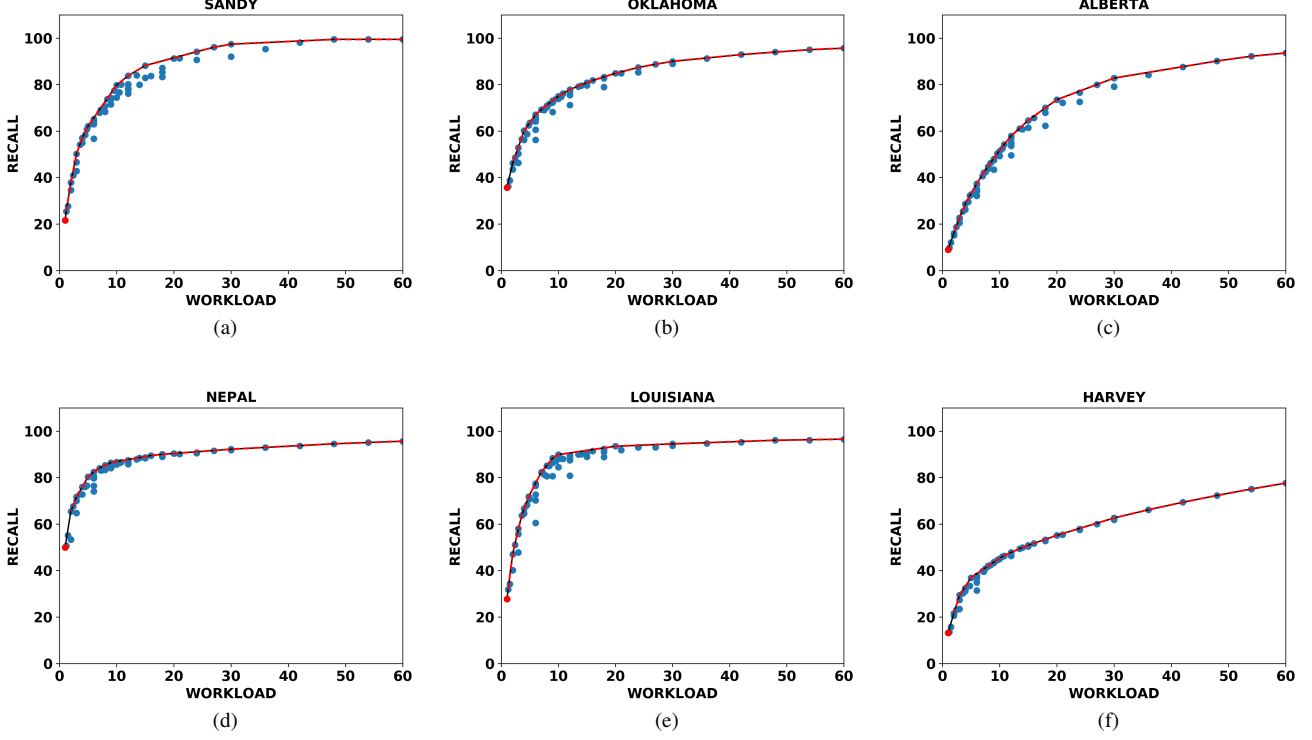


Fig. 1. Diminishing returns of a workload/attention budget. Red lines: maximum recall achievable given a budget. Blue dots: multiple average recall values for a given budget. [All higher resolution figures available at: <http://ist.gmu.edu/~hpurohit/informatics-lab/blogs/wi18-workload-aware-ranking.html>]

IV. EXPERIMENTS AND ANALYSIS

For a robust validation of our approach, we experiment with two schemes to generate RW matrices:

- *Periodic* algorithm processes messages posted in the time window of past H hours for generating top- k ranking and a RW matrix at the beginning of every hour (e.g., 7am, 8am). We consider $H = 24$ for a fair recommendation of recall and the required workload from a generated RW .
- *Realtime* algorithm processes messages posted in the time window of past G minutes for generating top- k ranking and a RW matrix at the beginning of every minute (e.g., 7:01am, 7:02am). We consider $G = 60$ for as accurate estimation as possible for the observed RW .

We borrowed the datasets of 6 crisis events (c.f. table III) from our prior work [7]. We processed them using the approach described in section III, within the periodic and realtime schemes. We used an existing epsilon-non-dominated

sorting algorithm for the Pareto optimization [30]. We analyze the following patterns for the relationship between the ranking performance and user workload achieved by the periodic and realtime schemes: patterns of recall versus workload recommendation and adaptive workload recommendation, and also, Pareto optimization comparison against the greedy baselines.

A. Recall vs Workload Recommendation Analysis

We studied the behavior of the average recall values for a value of workload and vice versa. In the highly stressful environment, the end users may not guarantee their availability for monitoring alerts consistently every hour. Thus, to support their decision making given such dynamic availability, the recommended RW matrix provides how much workload is necessary to achieve the desired recall (system performance) from the top- k alert ranking. Based on the periodic scheme, we computed the average values of recall and workload obtained across all the time slices of an event as shown in figure 1.

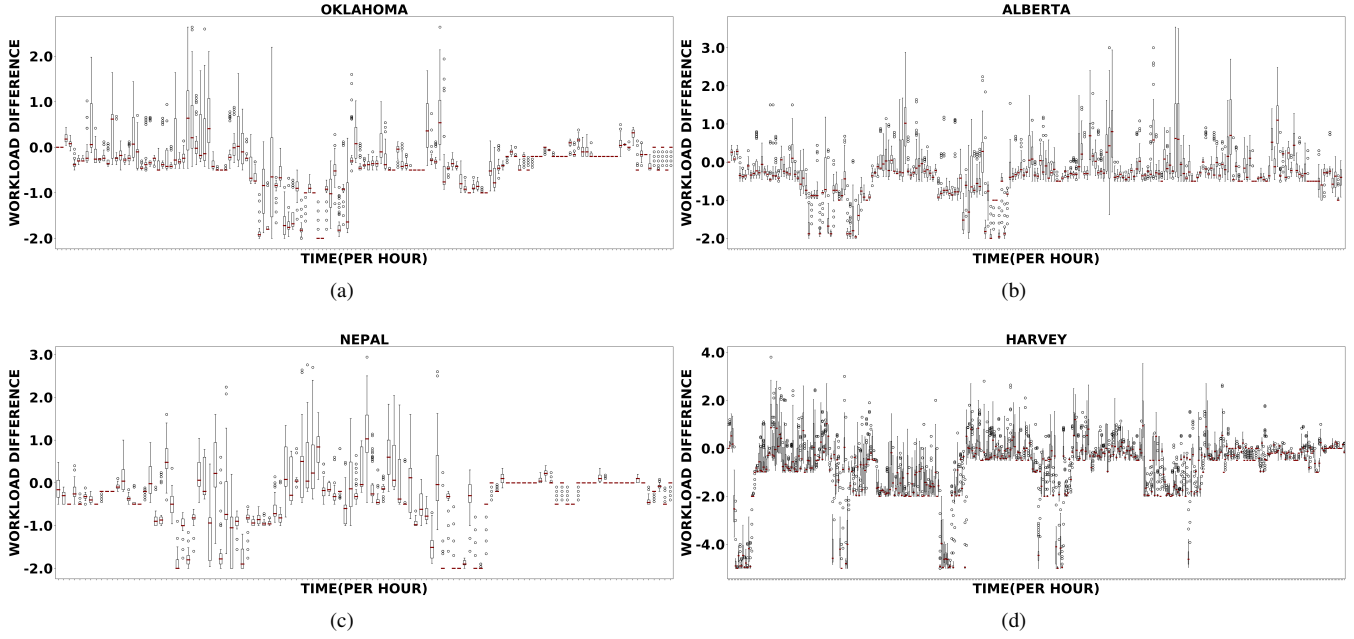


Fig. 2. Moving average difference between the minimum workload recommendations obtained by the periodic and realtime algorithms shows very small error, indicating the effectiveness of our periodic approach to recommend workload. [Due to space limitation we skipped the figures of two smallest events.]

Event-specific sub-figures demonstrate that there exists a pattern of multiple recall values (corresponding to different top- k rankings) for a given workload on x-axis, for instance, workload=10. The figures indicate the pattern of diminishing returns. The results also demonstrates the variability in the recall for the same workload across different events. Thus, one policy of selecting a specific top- k ranking and desired recall cannot be applied to all the events consistently.

B. Adaptive Workload Recommendation Errors

The analytical goal here is to analyze the error between the periodic predictive recommendations and the near realtime values of the required workload. We also assess the Pareto optimization performance against the baseline ranking selections.

1) *Error Analysis:* We first computed workload and recall for the hourly periodic algorithm output and then, the realtime algorithm output at every minute. We then measured the error difference between the outputs, followed by estimating the hourly mean and variance of error values. We observed that the error pattern was not contiguous for all the time slices across all the events. Therefore, we plotted the difference between the moving average values of each of the recall and workload metrics, where the average was computed across the sliding window of next 5 time periods. Figure 2 shows the pattern of a stable moving average for the error ranges across all the events, implying that the proposed periodic algorithm tends to rectify the estimation error in the recommendation in the near future. We further observed the bounded error ranges between the estimated and real values. The ranges are within 10% of the maximum possible workload (60), thus, showing the potential of the proposed approach to recommend the optimal values of workload and recalls for top- k alert rankings.

2) *Baseline Comparison – Greedy Selection:* We analyze the difference between the performance metrics obtained by our Pareto approach and two biased, greedy baselines.

First greedy approach relies on the policy of selecting the alert ranking with minimum workload recommendation every time and the second one relies on selecting the ranking with maximum recall recommendation. Figures 4 and 5 show the shortcoming of the greedy approaches where the choice of minimum feasible workload for recommendation does not always yield the maximum recall and thus, waste the time of the EOC personnel to review the irrelevant, useless messages.

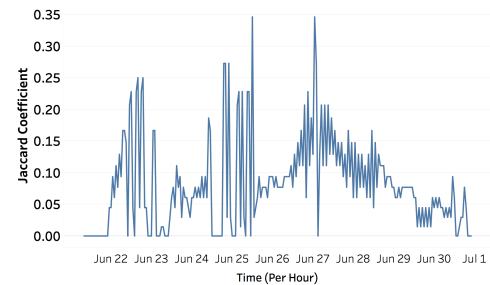
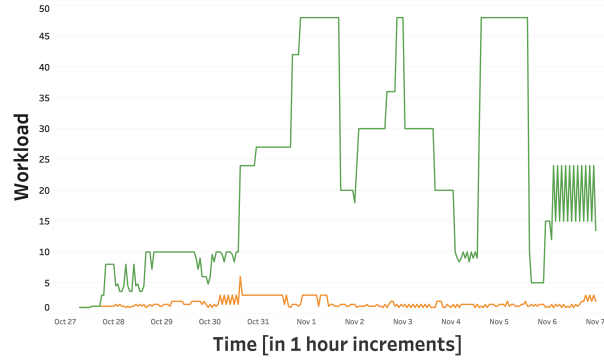


Fig. 3. Illustration of redundancy by the overlap of current top-50 alerts in the top-10 ranked alerts of the two succeeding time periods for Alberta event.

C. Redundancy and Timeliness

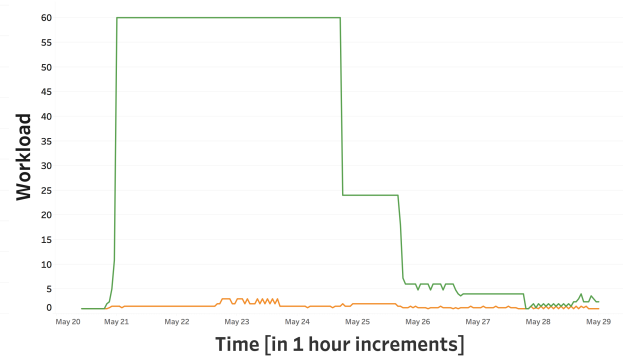
We also explored the information quality issues of the top- k ranked messages in a time period t_{ij} for redundancy. We computed redundancy using Jaccard Similarity between the set of top-10 alerts in t_{ij}^{s+1} and t_{ij}^{s+2} and the set of top-50 alerts in t_{ij}^s . Figure 3 shows the performance of periodic algorithm that re-surfaced the important messages as redundant alerts in future. This pattern suggests the need to efficiently factor redundancy and timeliness in the ranking computation.

Sandy: Greedy (Maximum Recall)



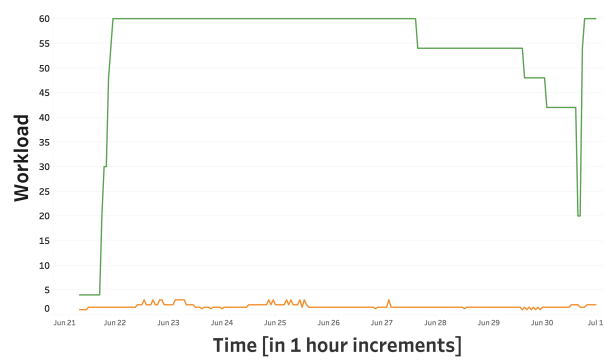
(a)

Oklahoma: Greedy (Maximum Recall)



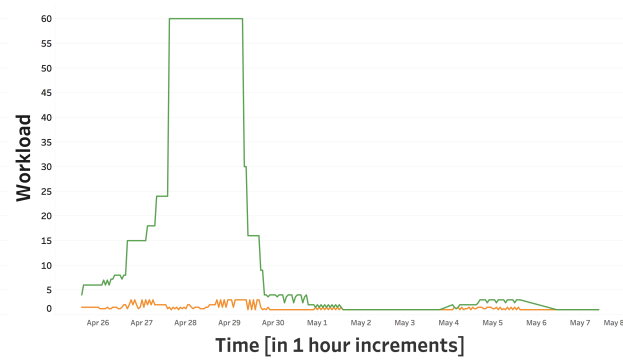
(b)

Alberta: Greedy (Maximum Recall)



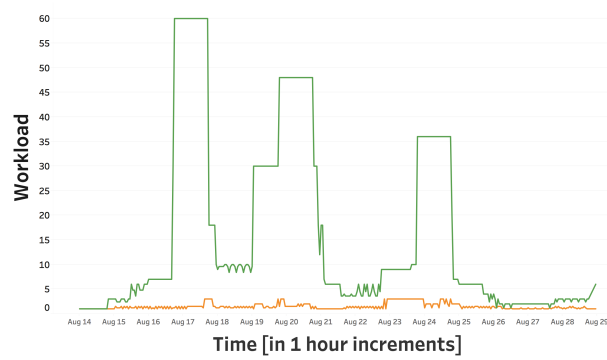
(c)

Nepal: Greedy (Maximum Recall)



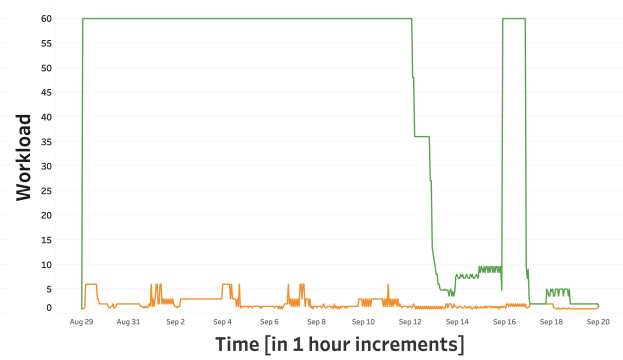
(d)

Louisiana: Greedy (Maximum Recall)



(e)

Harvey: Greedy (Maximum Recall)



(f)

Fig. 4. Lower value of minimum workload is recommended by the proposed Pareto algorithm (*orange*) in comparison to that obtained by the baseline, recall-based Greedy algorithm (*green*).

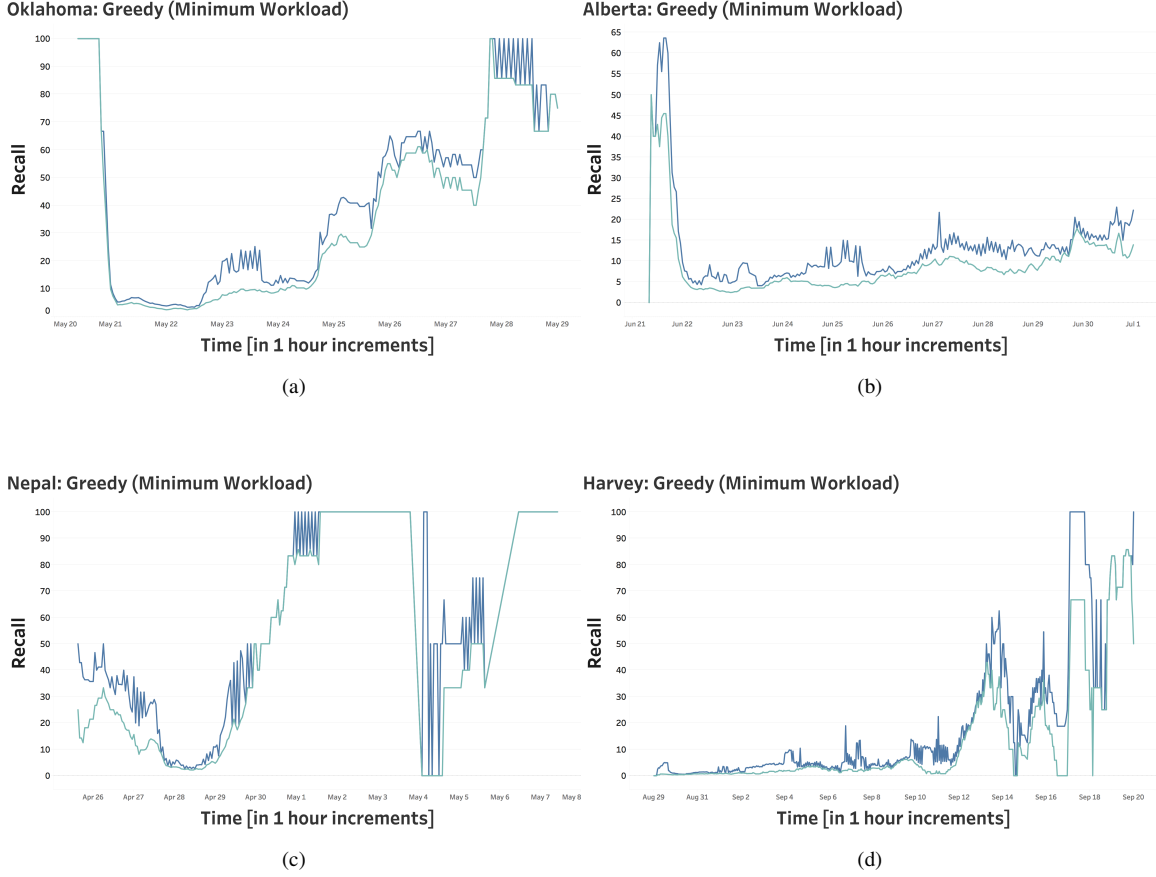


Fig. 5. Higher value of maximum recall at the minimum workload is attainable by the proposed Pareto algorithm (blue) in contrast to that obtained by the baseline, workload-based Greedy algorithm (green). [We skipped smallest two events due to space limitation.]

We also noticed that there are multiple choices in terms of which two points of k (for top- k) and t_{ij} (for period of computation) should be chosen. Table IV shows an illustration, if two points (k, t_{ij}) have the same workload, then we can create role-based, user-specific preference scheme. It is because every time the user receives an alert ranking list to review, s/he must switch work context. Therefore, we prefer to have lower frequency of requiring attention for user's review and recommend the smaller k and larger t_{ij} (less frequency to review without sacrificing performance).

V. LIMITATIONS

To the best of our knowledge this is a first study on the relationship between the performance metrics of alert ranking systems and the expected workload on end users in time-critical workplaces. Although this research serves as a preliminary work towards future research on user-aware adaptive ranking methods. Given the scope of our study, the additional analyses can be addressed as future work. First, we did not explore different types of ranking systems for our analysis, it is possible that alert-based and static IR-based ranking systems would perform differently. Second, we demonstrated our analyses using a specific domain's data, i.e. emergency

management, however, in other domains, the performance might vary in terms of the error between the estimated and real recommendations. Third, we have not studied the large range of workload bounds (considered only a reasonable range from 1 to 60 messages/hour) and its effects on the communications officers in EOC (e.g., one could hypothesize an excessive cognitive workload in the upper part of this range), which could be further studied. Lastly, the presented results depend on the relevance-based ranking system, however, there is a possibility to also incorporate redundancy and timeliness factors in the ranking system, which can be explored as a future work.

VI. CONCLUSIONS

Due to the limited budget of attention in the stressful environment of emergency management, traditional 'one-size-fits-all' solutions of alert generation for relevant social media updates are not effective. This paper presented a novel quantitative model for determining how many and how often should social media updates be generated, while also considering a given bound on the workload for an end user. Our formal model quantifies the relationship between the performance metrics of recall for top- k rankings and the required user

TABLE IV

ILLUSTRATION OF MULTIPLE CHOICES OF k AND t_{ij} RECOMMENDED FOR SANDY EVENT BY DIFFERENT RANKING SELECTION POLICIES, WHICH ALLOW USER-SPECIFIC PREFERENCE SCHEME. FOR INSTANCE, CONSIDER THE DESIRED RECALL OF 60% (**bold row**), A USER CAN CHOOSE SMALLER NUMBER OF k ALERTS AND LESS FREQUENCY OF t_{ij} BY THE PARETO APPROACH WITHOUT SACRIFICING PERFORMANCE.

Desired Recall	Recommended Workload	Pareto k	Pareto t_{ij}	Greedy-Recall k (baseline)	Greedy-Recall t_{ij} (baseline)	Greedy-Workload k (baseline)	Greedy-Workload t_{ij} (baseline)
10%	1.5	1	40	1	60	10	10
20%	3	1	20	5	60	10	10
..
40%	1	4	60	4	60	8	20
60%	1	1	60	3	60	4	60
..

workload. We presented an alert ranking system that employs a Pareto optimal algorithm for ranking selection, by adaptively determining the preference of top- k ranking and user workload over time. We presented empirical results based on real-world data from 6 crisis events to study the effects of different ranking selections and the trade-off with user workload, in comparison to different greedy baseline approaches. Our experiments demonstrate that the proposed approach can improve the efficiency of monitoring social media updates for EOC personnel while respecting constraints in user attention.

Reproducibility. Our dataset is available upon request, for research purposes.

VII. ACKNOWLEDGEMENT

Authors would like to thank reviewers for valuable feedback. Also, Purohit thanks US National Science Foundation grants IIS-1657379 & IIS-1815459 and Castillo thanks La Caixa project LCF/PR/PR16/11110009 for partial support.

REFERENCES

- [1] American Red Cross, "More americans using mobile apps in emergencies," August 2012, online and phone survey.
- [2] A. L. Hughes and L. Palen, "The evolving role of the public information officer: An examination of social media in emergency management," *Journal of Homeland Security and Emergency Management*, vol. 9, no. 1, 2012.
- [3] C. Castillo, *Big Crisis Data: Social Media in Disasters and Time-Critical Situations*. Cambridge University Press, 2016.
- [4] H. Purohit, C. Castillo, F. Diaz, A. Sheth, and P. Meier, "Emergency-relief coordination on social media: Automatically matching resource requests and offers," *First Monday*, vol. 19, no. 1, 2013.
- [5] X. He, D. Lu, D. Margolin, M. Wang, S. E. Idrissi, and Y.-R. Lin, "The signals and noise: Actionable information in improvised social media channels during a disaster," in *WebSci*, 2017, pp. 33–42.
- [6] K. Starbird, J. Maddock, M. Orand, P. Achterman, and R. M. Mason, "Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing," *iConference*, 2014.
- [7] H. Purohit, C. Castillo, M. Imran, and R. Pandey, "Social-eoc: Serviceability model to rank social media requests for emergency operation centers," in *ASONAM*, 2018, to appear. [Online]. Available: http://ist.gmu.edu/~hpurohit/informatics-lab/papers/serviceability_ranking_disasters_ASONAM18_final.pdf
- [8] M. Imran, C. Castillo, F. Diaz, and S. Vieweg, "Processing social media messages in mass emergency: A survey," *ACM Computing Surveys*, vol. 47, no. 4, p. 67, 2015.
- [9] M. Avvenuti, S. Cresci, A. Marchetti, C. Meletti, and M. Tesconi, "Ears (earthquake alert and report system): a real time decision support system for earthquake crisis management," in *KDD*, 2014, pp. 1749–1758.
- [10] J. Aslam, F. Diaz, M. Ekstrand-Abueg, R. McCreadie, V. Pavlu, and T. Sakai, "Trec 2014 temporal summarization track overview," NIST, Tech. Rep., 2015.
- [11] S. R. Hiltz, J. A. Kushma, and L. Plotnick, "Use of social media by us public sector emergency managers: Barriers and wish lists," in *ISCRAM*, 2014, pp. 602–611.
- [12] L. Plotnick, S. R. Hiltz, J. A. Kushma, and A. H. Tapia, "Red tape: Attitudes and issues related to use of social media by us county-level emergency managers," in *ISCRAM*, 2015, pp. 182–192.
- [13] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval: The Concepts and Technology Behind Search*, 2nd ed. USA: Addison-Wesley Publishing Company, 2008.
- [14] C. Rudin, "The p-norm push: A simple convex ranking algorithm that concentrates at the top of the list," *Journal of Machine Learning Research*, vol. 10, no. Oct, pp. 2233–2271, 2009.
- [15] L. Palen and K. M. Anderson, "Crisis informatics: new data for extraordinary times," *Science*, vol. 353, no. 6296, pp. 224–225, 2016.
- [16] S. M. W. G. U.S. Homeland Security Science & Technology, "Using social media for enhanced situational awareness and decision support," <https://www.dhs.gov/publication/using-social-media-enhanced-situational-awareness-decision-support>, 2014, accessed: 2018-06-12.
- [17] C. Reuter and T. Spielhofer, "Towards social resilience: A quantitative and qualitative survey on citizens' perception of social media in emergencies in europe," *Technological Forecasting and Social Change*, vol. 121, pp. 168–180, 2017.
- [18] T. Sakaki, M. Okazaki, and Y. Matsuo, "Tweet analysis for real-time event detection and earthquake reporting system development," *IEEE TKDE*, vol. 25, no. 4, pp. 919–931, 2013.
- [19] P. S. Earle, D. C. Bowden, and M. Guy, "Twitter earthquake detection: earthquake monitoring in a social world," *Annals of Geophysics*, vol. 54, no. 6, 2012.
- [20] B. Robinson, R. Power, and M. Cameron, "A sensitive twitter earthquake detector," in *WWW*, 2013, pp. 999–1002.
- [21] J. Yin, A. Lampert, M. Cameron, B. Robinson, and R. Power, "Using social media to enhance emergency situation awareness," *IEEE Intelligent Systems*, vol. 27, no. 6, pp. 52–59, 2012.
- [22] J. Bao and M. F. Mokbel, "Georank: an efficient location-aware news feed ranking system," in *GIS*, 2013, pp. 184–193.
- [23] C. Kedzie, K. McKeown, and F. Diaz, "Predicting salient updates for disaster summarization," in *ACL*, vol. 1, 2015, pp. 1608–1617.
- [24] R. McCreadie, C. Macdonald, and I. Ounis, "Incremental update summarization: Adaptive sentence selection based on prevalence and novelty," in *CIKM*, 2014, pp. 301–310.
- [25] K. Rudra, S. Ghosh, N. Ganguly, P. Goyal, and S. Ghosh, "Extracting situational information from microblogs during disaster events: a classification-summarization approach," in *CIKM*, 2015, pp. 583–592.
- [26] A. Nenkova and K. McKeown, "Automatic summarization," *Foundations and Trends® in Information Retrieval*, vol. 5, no. 2–3, 2011.
- [27] N. Sachdeva and P. Kumaraguru, "Call for service: Characterizing and modeling police response to serviceable requests on facebook," in *CSCW*, 2017, pp. 336–352.
- [28] T.-Y. Liu, "Learning to rank for information retrieval," *Foundations and Trends® in Information Retrieval*, vol. 3, no. 3, pp. 225–331, 2009.
- [29] S. A. Ross, "The economic theory of agency: The principal's problem," *The American Economic Review*, vol. 63, no. 2, pp. 134–139, 1973.
- [30] K. Deb, M. Mohan, and S. Mishra, "Evaluating the ϵ -domination based multi-objective evolutionary algorithm for a quick computation of pareto-optimal solutions," *Evolutionary computation*, vol. 13, no. 4, pp. 501–525, 2005.