Leveraging Non-Traditional Sources (NTS) for Space Situational Awareness (SSA) Analytics

Thomas R. Gemmer
Aptima, Inc.
Charlotte F. Shabarekh
Aptima, Inc.

ABSTRACT

Increasingly, eye witness observations about world and space events can be found in open source intelligence (OSINT) ranging from widely-used social media platforms such as Twitter to niche astronomy websites such as SeeSat. While much of Space Situational Awareness (SSA) continues to be performed using the United States Air Force's Space Surveillance Network (SSN), there is an increased interest in incorporating Non-Traditional Sources (NTS), such as OSINT and commercial sensors. However, working with NTS provides challenges for exploitation. Additionally, fusing and correlating NTS with traditional sensor data products is challenging because of inconsistencies in meta-data and reporting standards. For instance, backyard astronomers widely range in how they report observations and sometimes do not correlate observations to catalog IDs. Furthermore, exploiting NTS requires the consideration of the trustworthiness of sources.

In this paper, we fuse simulated observations from NTS with simulated observations from the SSN to improve our satellite maneuver prediction system which was presented at AMOS 2016 [1]. Our satellite maneuver prediction technology is built on an unsupervised machine learning algorithm called the Interval Similarity Model (ISM) which lends itself well for dealing with ambiguity and noisy data inherent in NTS. Its probabilistic approach can incorporate weights to award or penalize observations from specific sensors or data types and can in fact learn those biases to discover which sources are consistent with one another. We present experiments that compare the baseline system which only used SSN data with the new system that incorporates NTS and SSN observations. Our initial findings show that the updated algorithm improves better than the results in the 2016 paper and the probabilistic nature of this approach supports the bias or discounting of sources to manage issues of trustworthiness of NTS.

1. PROBLEM

Safe space operations begin with situational awareness, and we build space situational awareness (SSA) upon a foundation of accurate knowledge of the placement and activities of resident space objects (RSOs). The current Air Force space catalog includes over 17,000 actively tracked RSOs, with this number expected to grow by at least an order of magnitude later this year when the Air Force brings new Space Fence online [2]. Furthermore, there is a current push throughout the space community toward launching more, smaller satellites. The seemingly simple task of knowing where everything is in space is becoming more daunting every day. Advances in machine learning and automated RSO tracking technology are helping to improve SSA, however these analytics are dependent on the quality and cadence of the observations they are provided.

A promising mechanism to increase awareness is to incorporate observations from Non-Traditional Sources (NTS), such as the growing number of commercial, academic, and amateur operated sensors focused on space monitoring [3]. These sources could augment official Space Surveillance Network (SSN) observations to provide more comprehensive coverage of RSOs. However, there are a number of difficulties in using NTS, starting with the fact that there has traditionally been a rigorous vetting process, creating a high bar for use of a sensor's data for SSA. The primary issue is that NTS sensors that have not undergone any sort of qualification have unknown accuracy and precision. Incorporating such data blindly may not result in improved SSA, and may actually make the situation much worse by causing increased numbers of mistagged objects and uncorrelated track (UCT) detections. Related to that challenge, is managing inconsistent observations from NTS that conflict with observations from the SSN. Another issue is the lack of standardization of meta-data in non-traditional sources, compounding the already difficult problem of data fusion and integration. Making use of non-traditional sensor data will also require accounting for uncertain scheduling of observations since there is limited knowledge available of or influence over sensor tasking.

2. APPROACH

Leveraging NTS for SSA requires establishing trust in those sources for both the human operators and the analytic technologies that use them. Humans want to deliver robust inferences from information they trust to the decision-makers they support. Human analysts tend to judge source trustworthiness based on prior experiences and then verify and adjust their perception based on observations. When a source is truly unknown or incomparable with past experiences, the human's initial judgment tends to be driven by personal biases, leaving them more likely to reject that source in the future if an outcome is highly divergent from their expectations [4]. When conflicts between sources are identified, trust can potentially decrease across all sources, including the historically preferred sources, because even "trusted" sensors are susceptible to weathering events. For instance, while humans tend to establish trust with specific sources, they sometimes dramatically adjust their models of trust if a preferred source is unavailable or if a new source appears to be more trustworthy. The final model of how humans "stitch" information together into correlated, aggregated, and normalized knowledge is directly affected by their model of source trustworthiness. In other words, humans naturally strive derive trust from aggregate knowledge extracted over time, but can be heavily influenced, positively or negatively, by discrete observations from an individual source in the context of others.

In this paper, we explore the impact of incorporating data from NTS into our satellite maneuver prediction technology introduced at AMOS 2016 [1]. In a manner similar to the way humans establish trust in sources, our approach establishes trust in sources based on how well the observations contribute to the aggregate predictions, not the individual observations themselves. Our approach effectively rewards predictions that are consistent across sources and penalizes predictions that are inconsistent between sources.

Figure 1 shows the approach that we took to add NTS into our maneuver prediction technology. Our maneuver prediction algorithm works in two steps. First we learn a satellite's Pattern of Life (PoL), then we predict when it will most likely maneuver in the future. PoLs are repeatable, predictable behaviors that an object exhibits within a context and is driven by spatiotemporal, relational, environmental and physical constraints. An example of PoL are station-keeping maneuvers in GEO which become generally predictable as the satellite re-positions itself to account for orbital perturbations. These PoL are then used to create a Probability Density Function (PDF) for when a satellite is most likely to maneuver in the future. For the purposes of this paper, we will focus only on the maneuver prediction step.

To incorporate NTS, we use a two-step process that first performs maneuver prediction on each source individually and then compiles each source's prediction into a weighted, aggregate prediction. We begin to assign weights and confidence to a source by assessing its observations against historical maneuver data. This effectively begins tying the confidence measures directly to sources so that their reliability can be assessed over time. This approach has the benefit of modeling each source independently to allow for traceability of how each source contributed to the overall maneuver prediction. By splitting the data per source, we have fewer data points to create models from, however we mitigate the issue of data sparsity by combining the individual source's predictions into an overall, aggregated prediction in the second step. Machine learning tends to do better with more data and this is particularly true for unsupervised machine learning algorithms, such as we employ here (see Section 3 for algorithm details). To create the aggregated prediction, weights from each individual sensor are chosen to maximize source agreement. This is achieved by estimating model variance using the deviation from the most recently observed historical maneuver. The confidence measures which were initially assigned in the first step are updated with weights based on the sources overall contribution to the aggregate prediction.

3. ALGORITHM DETAILS

In this paper, we use the interval similarity model (ISM) with an updated workflow to establish trust in different input sources. The ISM uses a multi-step process iteratively calculates maneuver probabilities for multiple sources of observations and fuses those sources into a single combined probability.

3.1 Step 1

The ISM is an unsupervised machine learning algorithm which clusters temporal intervals based on periods of maneuvers and non-maneuvers. First step begins by calculating the probability that a satellite is executing a pattern

of maneuvers that are similar to an historical or ongoing PoL. Inspired by similarity-based clustering [5], the ISM's output is a probability density function (PDF) detailing the probability that a maneuver will occur with respect to time. Formally, suppose we have a given history of N ordered maneuver times: $\{t_1, t_N\}$. We use these historical maneuvers to build a PoL for the satellite consisting of the intervals between maneuvers, $\Delta t_{ij} = t_j - t_i$, as:

$$PoL = \{\delta_m = \Delta t_{kl}, \forall\, l > k\}$$

The ISM assumes that future maneuvers can be estimated from the distribution, P, of past maneuver intervals. Then the PoL represents a sample of P and we can use kernel density estimation (KDE) to approximate the density of the distribution, p, as [6]:

$$\hat{p}(t) = \frac{1}{mw} \sum_{n=1}^{m} K\left(\frac{t - \delta_n}{w}\right)$$

where the continuous variable t is the time from the latest maneuver (t_N), w is a smoothing parameter called the bandwidth, and $K(x)$ is a kernel function. The ISM uses a Gaussian based kernel:

$$K(x) \propto e^{-(x^2/2)}$$

on the assumption that errors in maneuver time values and future maneuver scheduling will be approximately Gaussian.

The ISM effectively assumes that each time interval, δ_m, it has seen before represents an interval that may be repeated. By assuming each prior interval is a Gaussian distributed representation of future maneuvers with width w, the ISM estimates the similarity between intervals by convolving all of these distributions so that similarly timed maneuvers become weighted more heavily. The choice of bandwidth (which is related to the standard deviation of the kernel), has a heavy effect on the shape of the final PDF and in this context, determines how closely spaced maneuver intervals must be to be considered similar. With a sufficient variety of data, this value could be learned to provide good values on different classes of satellites (e.g. ones that tend to maneuver daily versus weekly). For the purposes of this study, we fix w so that we can distinguish between maneuvers that occur more than approximately one hour apart.

3.2 Step 2

Data fusion between multiple sources is handled by running information from each source through the process outlined above, assuming some sources will detect and report nearly all maneuvers and some will catch very few. Each source reporting on the same RSO has its own PDF built up around the maneuvers that it knows about. Scoring of trust in each source is established by calculating the probability that the next maneuver should have occurred when it was observed by each source as:

$$S_m^i = \int_{\delta_{m+1}-\alpha}^{\delta_{m+1}+\alpha} \hat{p}^i(t)\, \mathrm{d}t$$

for each sensor, i, where α is a time range around the prediction that could be chosen operationally (lower values being associated with stringent requirements and lower scores overall). The scores are then normalized as:

$$\bar{S}^i = \sum_m S_m^i \left(\frac{1}{\sum_i \sum_m S_m^i}\right)$$

This normalized score is then multiplied by the PDF for each source, effectively scaling the amplitudes of the individual kernels from each source's observations. This allows the PDFs from each source to be convolved into a combined PDF, $\hat{p}_c(t)$, and maintain normalization on the final result.

$$\hat{p}_c(t) = \sum_i \bar{S}^i \, \hat{p}^i(t)$$

$$\int_0^{\delta_m} \hat{p}_c(t) \, \mathrm{d}t = 1$$

These steps are performed iteratively as new maneuvers are reported by each source so that source scores and predictions are updated in or near real time as new data becomes available. Figure 1 shows an example of the algorithm in operation sources 1-3 (blue, yellow, and green) have good agreement and source 4 (red) is quite different. Note that, while the out of phase source can exert some influence on the combined PDF, the three consistent sensors have a much greater effect.

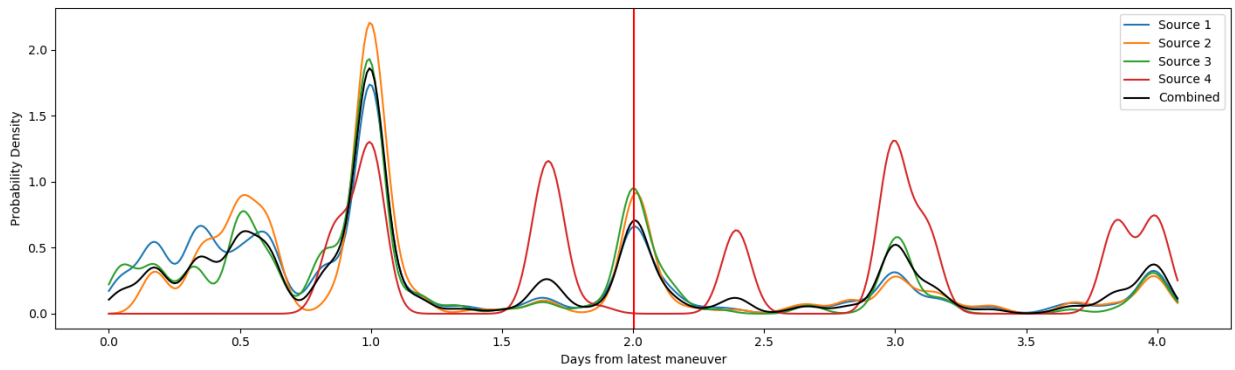

Figure 1: An example of the ISM in operation showing that while an out of phase source can exert some influence on the combined PDF (see peaks near 1.5 and 2.5), multiple consistent sensors have a much greater influence. The vertical, red bar indicates the actual time of the next maneuver.

4. RESULTS

The datasets used in this study consisted of maneuver times processed from simulated observations of Galaxy 15 (NORAD ID: 28884) and Anik F1R (NORAD ID: 28868) during a four-year period (2011-2015). The dataset contained synthetically generated astrometric data and demonstrated realistic levels of collection cadence (up to six days without observations of object) and noise (up to 90 µrad). In order to test source evaluation, we had each source report a subset of the complete maneuver list to the algorithm, with low reporting percentages representing NTS data and high percentages representing reliable sources (e.g. data derived from the SSN). This both keeps the scenarios realistic, since it is quite likely that especially NTS's would miss many maneuvers, and allows us to calculate a measure of how trustworthy each sensor should be since the fraction of maneuvers it reported is known.

Tables 1 and 2 show representative examples of source scoring results on each dataset under a number of different scenarios. The first three cases in each table demonstrate results for various numbers of sources each reporting some fraction of the actual maneuvers in each dataset. We calculate the expected trust score as a normalized fraction of the total maneuvers reported by each source, i.e. if M is the total number of maneuvers, and δM_i is the number of maneuvers reported by source i:

$$E(\bar{S}) = \frac{\delta M_i}{M} \left(\frac{1}{\sum_i \frac{\delta M_i}{M}} \right)$$

In all cases, the algorithm distributes trust among available sources consistently with the expectation.

In cases 1-3, all maneuvers being reported are true maneuvers. The final case examines a situation where source three is reporting random maneuver times. The algorithm suppresses trust on the random source in both datasets and redistributes it to the other sources. This shows that the algorithm is robust to very poor quality sources or sources that may be intentionally reporting incorrectly in an attempt to influence the results.

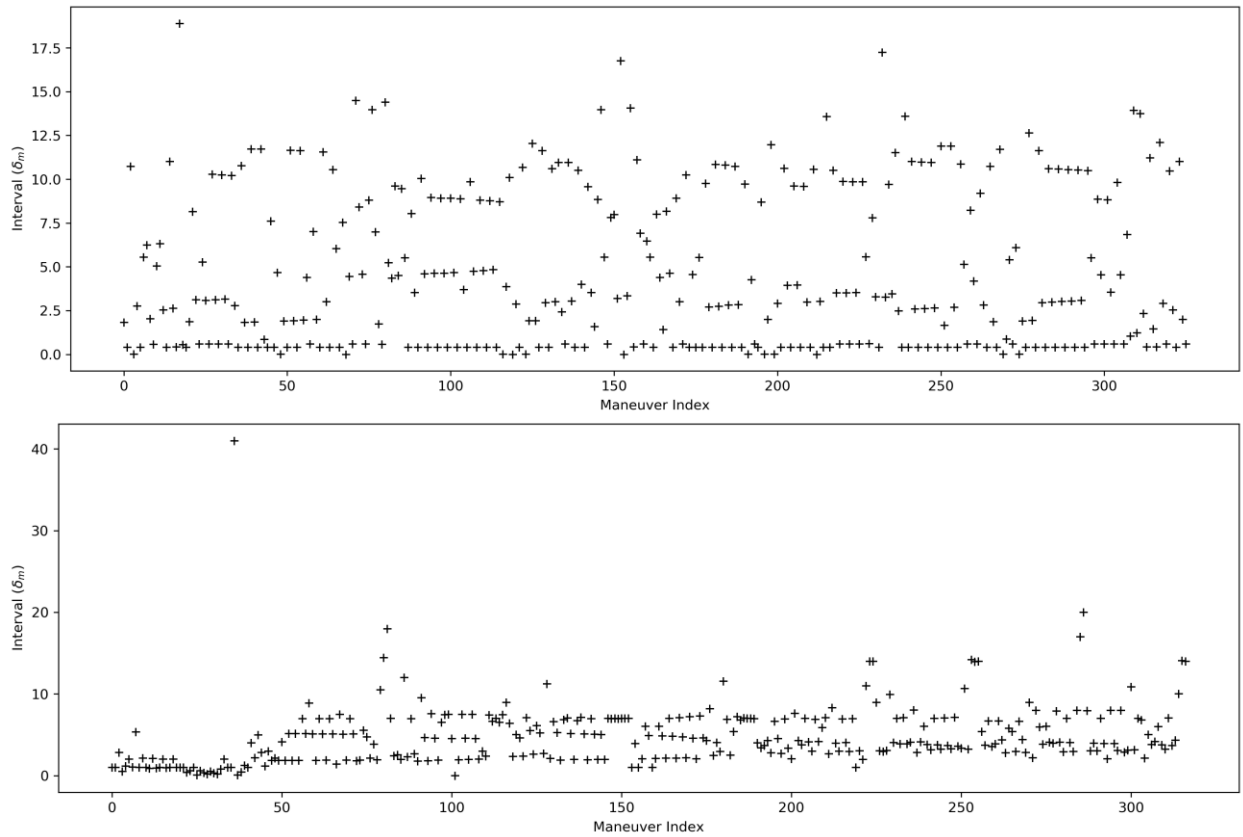

Figure 2: Depicting the intervals in each evaluation dataset (top – Anik F1R, bottom – Galaxy 15). Similar intervals can visually identified as horizontal groups. Note that both datasets contain periods of fairly regular maneuvers, periods of irregularity, as well as shifts in patterns that the ISM needs to track.

Table 1: Results from tests on Anik F1R dataset.

Case 1		Case 2		Case 3		Case 4	
Source (reporting percentage)	Trust Score (Expected)	Source (reporting percentage)	Trust Score (Expected)	Source (reporting percentage)	Trust Score (Expected)	Source (reporting percentage)	Trust Score
1 (100%)	0.375 (0.42)	1 (100%)	0.310 (0.36)	1 (100%)	0.278 (0.33)	1 (100%)	0.330
2 (80%)	0.316 (0.33)	2 (80%)	0.275 (0.29)	2 (80%)	0.245 (0.27)	2 (80%)	0.282
3 (60%)	0.309 (0.25)	3 (60%)	0.220 (0.21)	3 (60%)	0.208 (0.20)	3 (random)	0.047
		4 (40%)	0.195 (0.14)	4 (40%)	0.164 (0.13)	4 (40%)	0.180
				5 (20%)	0.106 (0.07)	5 (20%)	0.160

Table 2: Results from tests on Galaxy 15 dataset.

Case 1		Case 2		Case 3		Case 4	
Source (reporting percentage)	Trust Score (Expected)	Source (reporting percentage)	Trust Score (Expected)	Source (reporting percentage)	Trust Score (Expected)	Source (reporting percentage)	Trust Score
1 (100%)	0.376 (0.42)	1 (100%)	0.285 (0.36)	1 (100%)	0.267 (0.33)	1 (100%)	0.319
2 (80%)	0.328 (0.33)	2 (80%)	0.274 (0.29)	2 (80%)	0.245 (0.27)	2 (80%)	0.287
3 (60%)	0.297 (0.25)	3 (60%)	0.236 (0.21)	3 (60%)	0.219 (0.20)	3 (random)	0.039
		4 (40%)	0.206 (0.14)	4 (40%)	0.183 (0.13)	4 (40%)	0.230
				5 (20%)	0.086 (0.07)	5 (20%)	0.125

We define the overall quality of the ISM predictions in two ways, both examining how well the combined PDF predicted the actual next maneuver in each iteration. The first is to examine the total predicted probability of the next maneuver occurring within +/- 12 hours of its actual time, p_{12}:

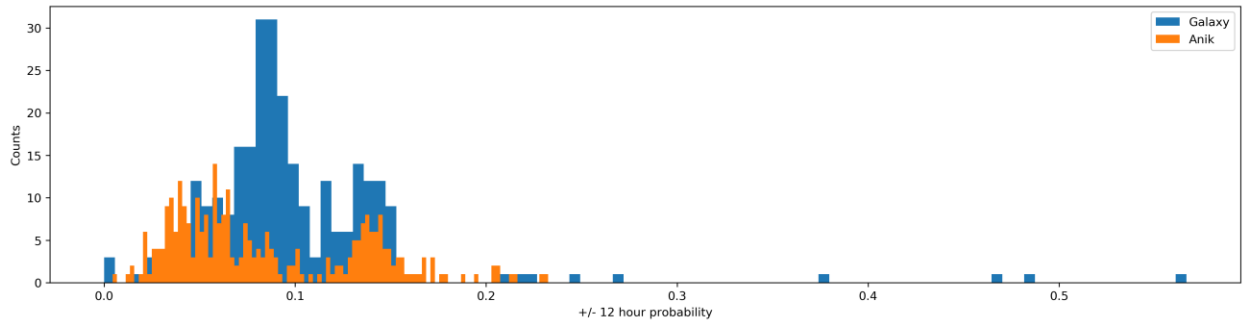

Figure 3: Histogram for p_{12} on each dataset.

$$p_{12} = \int_{\delta_{m+1}-12}^{\delta_{m+1}+12} \hat{p}_c(t)\, \mathrm{d}t$$

Histograms of this parameter for each dataset are shown in figure 3. The bimodality of the distribution in both cases is an interesting feature of note. This is caused by the fact that both datasets have intermittent periods of highly regular maneuver intervals (see figure 2). During those regular periods, the combined PDF develops one to three strong peaks, tending to drive the cumulative probabilities up.

The second parameter used is the difference in time between a new maneuver and the closest strong peak in the combined PDF. A strong peak being defined as any peak with a density over 20% of the largest peak (in both datasets, there is an average of one such peak for every two days in the prediction window). The histograms for this parameter on each dataset are shown in figure 4. Median values are 3.4 hours for Galaxy 15 and 3.8 hours for Anik F1R, and the 90th percentiles occur at 12.4 and 23.3 hours respectively.

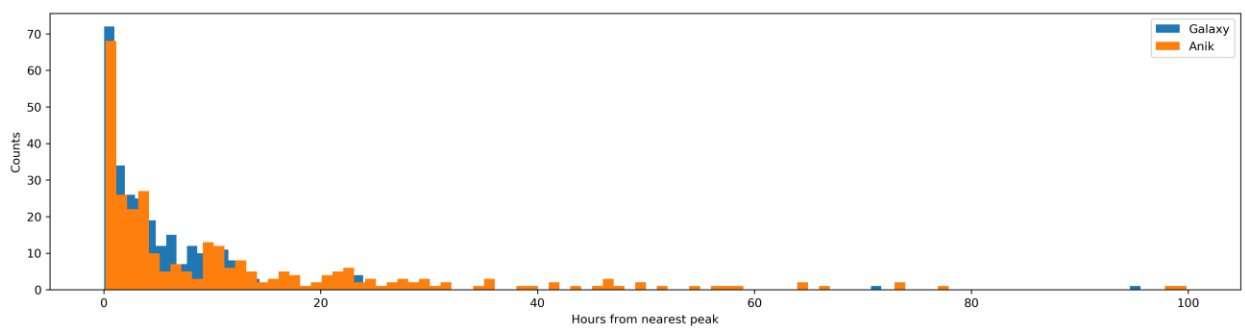

Figure 4: Histogram of the difference in time between a new maneuver and the closest strong peak in the combined PDF on each dataset.

5. CONCLUSIONS

The interval similarity method outlined in this work can be used to increase awareness of unanticipated dynamic events to provide operators with maximum time to generate and execute courses of action based on multiple data sources of uncertain trustworthiness. The ISM outlined here has demonstrated the ability to consume RSO maneuver information from multiple sources with variable reliability and autonomously assess the trustworthiness of the sources consistently with expectations. We have additionally demonstrated the method's ability to detect 90% of maneuvers correctly to within 12-24 hours of their actual time and 50% of maneuvers to within 3-4 hours of their actual time on realistic datasets. The ISM works in real time and can be extended to an arbitrary number of sources.

An additional important feature of this method is the maintenance of traceability for its confidence assessments from final scores and probability assessments back to the source level through the individual source PDFs. Both raw and normalized scores can also be tracked over time to allow an operator to have insight into how data quality is evolving overall and from each source. It is also important to note that the data fusion step of the ISM in particular is not restricted to this use case and can be applied in any situation where event probabilities are being estimated by PDFs using data from multiple sources (e.g. conjunction assessments).

6. REFERENCES

[1] C. Shabarekh, J. Kent-Bryant, and G. Keselman, "A Novel Method for Satellite Maneuver Prediction," in *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS)*, 2016.

[2] J. A. Haimerl and G. P. Fonder, "Space Fence System Overview," in *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS)*, 2015.

[3] L. C. J. Raley, R. M. Weisman, C. C. Chow, M. Czajkowski, and K. Sotzen, "OrbitOutlook: Autonomous Verification and Validation of Non-Traditional Data for Improved Space Situational Awareness," in *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS)*, 2016.

[4] A. Fouse, R. S. Mullins, G. Ganberg, and C. Weiss, "The Evolution of User Experiences and Interfaces for Delivering Context-Aware Recommendations to Information Analysts," in *Advances in Intelligent Systems and Computing*, 1st ed., T. Ahram and C. Falcão, Eds. Springer, Cham, 2018, pp. 15–26.

[5] M.-F. Balcan, A. Blum, and S. Vempala, "Clustering via Similarity Functions: Theoretical Foundations and Algorithms," in *Proceedings of the 40th ACM Symposium on Theory of Computing (STOC)*, 2008.

[6] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, 1st ed. New York: Routledge, 2018.