Optimizing Network Performance, Automation, and Intelligent Decision-Making through Real-Time Big Data Analytics

Ahmed Hassan
Department of Information Systems, University of Babylon,

Ali H. Mhmood
Tikrit University

Abstract

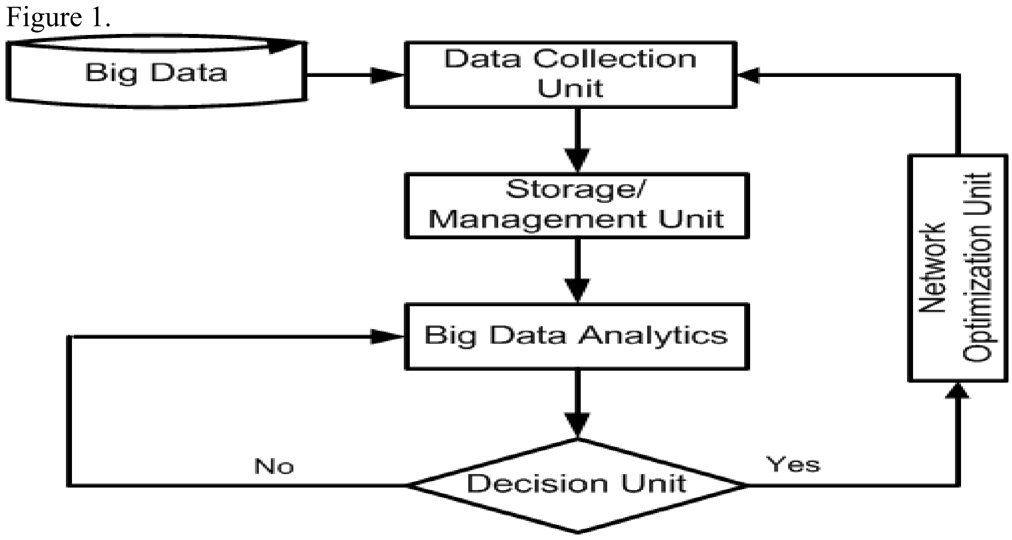
Network performance optimization and automation are critical for providing quality digital services and enabling data-driven decision making. This paper examines how real-time big data analytics can be leveraged to optimize network operations, enable intelligent automation, and empower data-driven decision making. A conceptual framework is presented illustrating the key components for building a real-time analytics solution including data acquisition, stream processing, data warehousing, complex event processing, predictive modeling, visualization, and automation. Critical technical and organizational enablers are discussed. Through an extensive literature review and real-world use cases, the paper demonstrates the application of real-time analytics optimizing performance across network domains including capacity planning, traffic engineering, cybersecurity, customer experience management and network operations automation. Challenges in adopting real-time analytics are analyzed and mitigation strategies proposed. Overall, the paper highlights the immense potential of real-time big data analytics in driving the next phase of innovation for communication service providers by enabling intelligent, automated, and optimized network performance along with data-empowered decision making.

Indexing terms: Bike sharing programs, Sustainable transportation, Modal shift, Last-mile connectivity, Data-driven planning

Introduction

The advent of 5G and proliferation of smart devices is fueling an explosion of data traffic on communication service provider (CSP) networks. Emerging digital services like autonomous vehicles, industrial automation, remote healthcare and metaverse will further drive the data deluge. To keep pace, CSPs are transforming their network and service architectures by adopting cloud native technologies like network function virtualization (NFV), software defined networking (SDN) and multi-access edge computing (MEC). While these disaggregated and software-driven architectures deliver flexibility and agility, they also greatly increase complexity and dynamics of networks [1]. Moreover, with shifting competition and new data-driven digital services, CSPs need deeper real-time network insights to make strategic decisions on 5G investments, infrastructure expansion, and service innovations.

Figure 1.



To cope with rapidly evolving networks and deliver hyper-automated and intelligent connectivity, CSPs need to harness the power of real-time big data analytics. With streaming network data processed through advanced analytics and machine learning algorithms, operators can gain end-to-end visibility, optimize and automate network operations, and enable data-driven decision making. Real-time analytics delivers key advantages over traditional daily or hourly batch processing of network data by enabling instantaneous insights and triggers for automated actions. This empowers network automation and intelligent actions not previously possible with latency of offline analytics [2], [3].However, building real-time analytics on telco networks involves surmounting significant technology and organization challenges due to the velocity, variety, volume and veracity of data. From the technology perspective, highly scalable and resilient platforms are needed for acquiring, storing and analyzing the endless streams of network data. Organizationally, new roles, processes and cultural shifts are required to leverage analytics for network planning and operations. Despite these barriers, early adopters of real-time telco analytics are already achieving tremendous benefits including up to 25% optimization in capital expenses, 55% reduction in network faults through predictive maintenance and three-fold faster resolution of customer complaints via real-time dashboards [4].

Therefore, this paper examines the application of real-time big data analytics in optimizing CSP network operations and empowering data-driven decision making. The conceptual foundations and key enablers are presented. Through an extensive literature review and real-world use cases, the functional areas where real-time analytics unlocks value are analyzed including network capacity planning, service assurance, cybersecurity, customer experience management and network operations automation. Finally, adoption challenges are discussed along with mitigation strategies [5].

Conceptual Framework

The key components in a real-time telco analytics solution include data acquisition, stream processing, data warehousing, analytics, modeling, visualization and automation. These enable a closed-loop system where analytics models guide automation and vice versa.

Data Acquisition: The first vital step is acquiring high velocity data streams from the myriad network domains into an enterprise data bus or lake serving as the raw material for analytics. This requires consolidating and collecting data from diverse sources spanning network equipment, virtualized functions, operational support systems as well as customer-facing applications [6]. Standard interfaces like NETCONF, RESTCONF and gRPC provide abstract and open APIs to extract streaming telemetry and event data from physical and virtual network functions. Intelligent network interface cards and routers embed efficient streaming with enhanced telemetry collection mechanisms like INT, IOAM and IFA defining structured data formats for simplified analytics [7]. Legacy network devices require additional probes and adapters to capture traffic statistics, utilization metrics, routing changes and failure events.

Other key data sources like billing records, customer tickets, network inventory databases and data warehouses need custom built interfaces and extractors to acquire daily and hourly streams. A common data endpoint or message bus enables consolidation of data from disparate interfaces into a unified data foundation. The streaming acquisition process applies schema validation, security rules, tagging sensitive fields, anonymization, geospatial and temporal metadata enrichment to prepare raw data for analysis systems. Overall an enterprise data streaming pipeline transports telco data streams from diverse network and business functions into common data lake repository tiered across hot, warm and cold storage for cost efficiency. The data lake approach provides flexibility to accommodate varied data types and schema changes typical for rapidly evolving virtualized networks.

Stream Processing: High velocity telco data after acquisition flows into scalable distributed stream processing engines for filtering, aggregation and simple rule-based analysis before further analytics. This is necessitated by the need to handle millions of records and events per second demanding scale-out software frameworks leveraging clustered commodity infrastructure over proprietary stream computing hardware. Widely used distributed stream processing platforms include open source Apache Spark Streaming, Flink, Storm and Kafka Streams while cloud hosted options offer managed turnkey solutions. Core processing techniques apply windowing constructs over streams to define duration based bounds for aggregation [8]. Operations like data cleansing and normalization apply common transformations for downstream cohesion. Filters allows selective routing of data to multiple outputs based on conditions. Stream processing provides vital data preparation offloading initial heavy lifting from subsequent analytics stages.

Table 1: Key Data Sources for Real-Time Telco Analytics

Data Source	Type	Analytics Use Cases
Call Detail Records	Structured Stream	Revenue Assurance, Fraud Detection, Customer Churn
Network Traffic Data	Time Series	Capacity Planning, Congestion Management
Alarms and Events	Time Series of Events	Service Assurance, Root Cause Analysis
Customer Interactions	Text Streams	Sentiment Analysis, Recommendations
Network Inventory	Relational DB	Topology Analysis and Visualization
APIs and Messages	JSON Streams	Orchestration Analytics

Basic event pattern analysis and simple rule-based decisions can also be executed for achieving ultra-low latency use cases like fraud detection. However advanced analytics spans predictive models and machine learning outside the scope of stream processing capabilities demanding additional layers [9]. Processed stream outputs are delivered to multiple end points including short term storage for retention ranging few hours for real-time monitoring to few weeks for detailed analysis. Selected data like call records get ingested into cloud data warehouses for long term big data analytics. Transactional data updates enterprise databases and data marts powering customer-facing applications and OSS/BSS systems.

Data Warehousing: While streaming and complex event analytics provides transient analysis on data in motion, historical big data analytics relies on structured storage of network traffic details, call records, customer events and inventory databases. This analytical data warehousing leverages distributed clustered database frameworks optimized for adhoc complex querying spanning terabytes to petabytes from enterprises data lakes [10]. Managed cloud data warehouse solutions include Snowflake, BigQuery, Redshift providing easy scale-out on cheap cloud object storage. Sophisticated caching, indexing, compression, partitioning techniques enable interactive analysis via SQL interfaces performant even on billions of records. Historical data combined with real-time streams power advanced time series forecasting algorithms crucial for long range predictive analytics [11].Data warehouses follow systematic ingestion workflows transforming selected raw data streams using schema definitions and mapping rules before loading into various database schemas arranged for efficient analytic modeling. Regular data archiving to cheaper tiers and expiration polices based on usage patterns helps optimize storage costs without affecting analytic value.

Complex Event Processing: Alongside historical big data warehousing, streaming analytics hinges on complex event processing (CEP) whereinference rules instantly analyze events on arrival for actionable alerts and trigger automation workflows without persistence. For instance, KPI threshold violations can activate critical alerts while usage anomalies may indicate security threats or network faults for investigation. Distributed CEP frameworks like Apache Flink apply sophisticated pattern matching traversing thousands of rules to uncover situational insights from massive event streams not needing offline storage [12]. The ephemeral in-memory event pattern analysis handles data velocities and transiency not suitable for HDFS and database infrastructure. This empowers real-time automation use cases otherwise not possible with data warehouse processing constraints [13].

Table 2: Machine Learning Models for Real-Time Telco Analytics

ML Model	Algorithms	Analytics Applications
Classification	Logistic Regression, SVM, Neural Nets	Anomaly Detection, Predictive Maintenance
Forecasting	ARIMA, RNN	Network Traffic Prediction

Recommendation	Collaborative Filtering, Content-based Filtering	Personalized Customer Offers
Natural Language Processing	RNN, BERT, Transformers	Speech Analytics, Social Media Monitoring
Reinforcement Learning	Q-Learning, SARSA, Deep Q-Networks	Resource Optimization, Cognitive Management

Predictive Modeling: Machine learning and predictive analytics entails training models using supervised, unsupervised and reinforcement learning algorithms over historical and real-time data to uncover hidden insights that can forecast impending failures, future demands and optimal actions. The trained models once validated and optimized then score streaming data from network devices and applications for intelligent run-time decisions powering automation systems [14]. ML models like classification identify anomalies in traffic, classify network threats, predict churn etc while natural language techniques empower speech analytics and social media analytics. Forecasting models enable data-driven capacity planning matching dynamic user demands. Recommender systems guide personalized customer promotions and care. Reinforcement learning optimizes intricate network configurations and cloud resource adaptations as learned over data for always-on improvements.

Predictive model development requires specialized data science platforms supporting iterative workflows spanning pre-processing, feature engineering, model prototyping, parameter tuning and statistical evaluation. Model development may leverage notebooks, open-source libraries like Tensorflow and cloud machine learning toolkits. The published ML models integrate with streaming pipelines and automation engines for low latency scoring of network events and telemetry unpacked for decisions.

Data Visualization and Reporting: Actionable insights necessitate intuitive visualization allowing all personas from network technicians to C-suite executives to rapidly comprehend massive volumes of data streams through smart graphical dashboards organized for operational decision making. Integrated developer toolkits simplify building customizable real-time dashboards tuned to domain specific monitoring and diagnostic needs. Network engineers rely on graphical topology views with node and link KPI overlays to instantly isolate fault domains from massive device level metrics streaming at sub minute intervals. Customer support teams leverage conversational analysis dashboards with cohort filters revealing rising complaints to prioritize. Executives prefer holistic network health dashboards with geography based heatmaps to track quality of experience.

Carrier grade visualization technology supports category-based correlation analysis with case management, automated anomaly detection and collaborative workflows for accelerated problem resolution. Geospatial techniques help uncover geographic patterns leveraging network inventory integration. Outlier analysis identifies deviations prompting further investigation. Automated reporting distributes insights to stakeholders while self service capabilities enable user defined analyses.

Network Automation: A strategic benefit of real-time telco analytics lies in driving extensive automation across network management processes spanning demand planning to predictive maintenance and service rollouts thereby limiting repetitive manual efforts allowing human creativity and innovation. This necessitates tight confluence between analytics systems and software-driven infrastructure enabled by contemporary SDN and NFV technology foundations with open interfaces and a flexible control plane. Analytic model scoring combined with declarative intent policies get interpreted by SDN controllers automatically pushing network-wide configurations and virtual function deployments matching the desired state [15]. Likewise streaming anomalies from ML models trigger automated runbooks launching workflows handing off provisioning tasks through workload orchestrators and cloud management platforms activating self-healing functions. Chatbots empower helpdesk automation for intelligent subscriber assistance. Recommender systems enable targeted campaign execution systems triggering personalized customer nudges through interaction channels.

Enablers: Realizing the conceptual framework needs strong organizational sponsorship and multi-disciplinary teams spanning data engineers, scientists, automation experts and domain specialists. Agile pilots demonstrating quick wins build further support and feed lessons learned. As real-time data platforms support mission critical networks, high availability design and leveraging cloud native architectures promotes resilience and scalability. Mature data governance managing security, quality, lifecycle and access policies cultivate trust and adoption. Furthermore, real-time analytics imposes massive data storage and processing needs demanding optimization of cloud costs using tiered storage, auto-scaling and spot instances. Overall, real-time telco analytics is a continuous journey needing sustained commitment to maximizing analytic value.

Optimizing Network Planning with Real-Time Analytics

Network capacity planning forms the foundation on which high-quality customer experience and new digital services depend. Conventionally confined to periodic efforts forecasting long-term capacity expansions, the dynamism introduced by 5G and virtualization necessitates continuous and data-driven network optimization based on usage insights from streaming analytics.

The following use cases demonstrate the role of real-time analytics in network planning.

Intelligent Spectrum Management: Allocating radio access network (RAN) spectrum to meet the escalating demands of diverse services, such as IoT, mobile broadband, and critical applications, poses a considerable challenge in the face of fluctuating traffic patterns. Real-time analytics emerges as a pivotal solution, offering a nuanced understanding of application consumption, device distribution, and geographical traffic dynamics. This granular visibility facilitates dynamic spectrum balancing, allowing network operators to optimize resource allocation in response to the evolving needs of different services [16]. Moreover, the integration of traffic predictive models further enhances the efficacy of real-time analytics by enabling the identification of congestion hotspots. This predictive capability empowers operators to implement targeted augmentation strategies, ensuring that network resources are efficiently deployed where they are most needed, thereby enhancing overall network performance and user experience. In essence, the application of real-time analytics in spectrum allocation not only addresses the complexities of contemporary service demands but also contributes to the proactive management of network congestion, fostering a more responsive and resilient radio access network.

Augmenting Microwave Backhaul: Growing end-user bandwidth strains existing microwave backhaul necessitating costly fiber buildouts. Streaming analytics helps determine precise capacity shortfalls across backhaul links based on live traffic loads. This allows targeted microwave modernization while optimizing fiber investment.

Informed Densification and Cloud Migration: Densifying metro networks with mini-data centers and central offices requires forecasting user demand, application types, edge resource needs and cloud migrations. Real-time analytics provides data-driven projections for optimal densification planning minimizing overprovisioning. The above use cases highlight the role of real-time analytics in enabling data-driven planning reducing guesswork and stranded capacity from periodic estimation exercises. Stream analytics delivers granular and live demand visibility for precise and dynamic network augmentation aligned to users. Next, we examine how real-time analytics transforms network operations.

Transforming Service Assurance with Real-Time Analytics: Delivering service availability and performance as per SLAs across dynamically changing virtual networks demands real-time assurance capabilities harnessing analytics and automation. Batch processing delays hindering rapid diagnosis and lowering mean time to resolution motivates streaming analytics-based assurance. Key use cases follow.

Predictive Network Maintenance: Complex virtualized networks suffer dynamic faults from NFV bugs, complex dependencies and misconfigurations. Real-time analytics leveraging ML predictive models identify probable outages from precursors in streaming KPI metrics, telemetry events, virtual network function (VNF) logs and trouble tickets. This enables fix-before-fail proactive maintenance boosting availability.

Rapid Root Cause Isolation: When outages inevitably occur, real-time topology visualization, automated traversing and rule/ML-based event correlation quickly pinpoints root causes by analyzing streaming VM, container logs, infrastructure metrics and VNF events. Near real-time monitoring analytics reduces diagnosis from hours to minutes [17].

Smart Network and IT Resource Optimization: Inefficient network and IT resource usage strains availability and costs. Analyzing streaming container orchestrator events using vectorized anomaly detection reveals oversized VM allocations for rightsizing. Streaming metrics from hosts combined with smart load balancing optimizes utilization. The above intelligent assurance use cases powered by real-time analytics and ML deliver higher automation, lower faults and optimal utilization even as virtual networks grow more complex. Next, we examine security applications.

Table 3: Key Cultural Shifts for Real-Time Analytics Adoption

Dimension	Traditional Culture	Analytics Culture
Decision Making	Intuition and Periodic Analysis	Continuous Data-driven Decisions
Network Planning	Manual Capacity Forecasting	Automated Traffic Prediction-based Planning
Service Operations	Reactive Troubleshooting	Predictive and Proactive Optimization
Customer Interactions	Rule-based Engagement	Personalized Context-aware Engagement
Network Management	Expert System Guided	Continual Machine Learning based

Enhancing Cybersecurity with Real-Time Analytics

While connectivity revenues decline, cyber threats explode requiring CSPs continuously harden network infrastructure and services. ML-driven real-time analytics boosts threat visibility, detection and response capabilities across the expansive attack surface.

Holistic Security Intelligence: The proliferation of numerous monitoring tools in organizational cybersecurity landscapes often results in the creation of isolated security data silos. This fragmentation poses a significant challenge to achieving comprehensive threat awareness across the entire organization, impeding the ability to mount coordinated responses to potential security incidents. To address this issue, the implementation of streaming normalization, correlation, and aggregation through big data analytics emerges as a pivotal solution. By harnessing the power of these analytical techniques, disparate security events from various sources are harmonized into a unified and coherent format. This unified visibility enables security professionals to gain a holistic understanding of the threat landscape [18]. Furthermore, the integration of risk scoring mechanisms facilitates the prioritization of security incidents based on their potential impact, allowing organizations to allocate resources efficiently and respond promptly to the most critical threats. In essence, the adoption of big data analytics in security operations not only breaks down the silos that impede threat awareness but also empowers organizations to enhance their overall cybersecurity posture through strategic and informed decision-making [19].

Rapid Threat Detection and Containment: Traditional intrusion detection systems (IDS) built on rule-based methodologies often grapple with the challenge of producing an overwhelming number of false alerts, thereby inundating security teams and hindering their ability to discern genuine threats from noise. Moreover, these conventional systems may exhibit limitations in detecting highly sophisticated threats, such as novel zero-day attacks, which operate outside the parameters defined by pre-existing rules. To address these shortcomings, a paradigm shift toward streaming anomaly detection, empowered by unsupervised machine learning (ML), has gained prominence in contemporary cybersecurity strategies.

Proactive Infrastructure Vulnerability Management: Virtual infrastructure sprawl makes securing vulnerabilities before exploitation challenging. Streaming analytics on container registry events, vulnerability advisories and patch releases combined with ML-based risk scoring initiates automated patching of critical vulnerabilities. The presented cybersecurity use cases demonstrate how real-time analytics addresses blindspots from fragmented security monitoring while enabling intelligent threat detection/response closing attack exposure gaps. Next, customer experience applications are discussed.

Optimizing Customer Experience with Real-Time Analytics

As 5G opens opportunities for innovative vertical use cases and building deeper engagement, analytics-driven customer experience management becomes pivotal for retention and growth. Key solution areas follow.

Holistic Customer Intelligence: Fragmented IT systems pose a significant challenge in the realm of customer interactions by scattering data across multiple channels, thereby creating visibility blind spots and hindering context awareness. The lack of a unified view makes it difficult to understand and respond to customer needs effectively.

Streaming analytics emerges as a pivotal solution to this predicament by reconstructing a comprehensive profile that encompasses a customer's likes, dislikes, usage patterns, interactions, and relevant events. This holistic view enables organizations to achieve personalized and real-time engagement with their customers. By leveraging streaming analytics, businesses can break down silos, seamlessly integrate data from diverse sources, and gain a nuanced understanding of customer behavior [20]. This, in turn, empowers enterprises to deliver tailored and contextually relevant experiences, fostering customer satisfaction and loyalty in an increasingly dynamic and competitive market.

Augmented Customer Care: Long call routing and transfers result in tedious service journeys. Real-time speech analytics of customer calls linked with account history automatically retrieves profiles for hyper-personalized attention slashing efforts. Sentiment detection guides agent responses.

Contextual Upsell and Cross-sell recommendations: Periodic offline analytics lacks consumable insights to capitalize micro-moments. Event stream processing harnessing usage context, augmented by AI recommenders dynamically proposes upsell and cross-sell offerings creating value. The above streaming analytics use cases help transform static customer relationships into intelligent life-cycle partnerships via continuous context-aware servicing helping capture micro opportunities. Finally, we analyze automation use cases.

Driving Network Automation with Real-Time Analytics

Dynamic network changes coupled with rising complexities make complete human-driven operations infeasible necessitating closed-loop automation. Analytics and automation must tightly coalesce for functional agility as elaborated below.

Intent-driven Orchestration: Manual network provisioning, with its inherent limitations, often finds itself grappling to align with the dynamic and volatile intents of users. The conventional approach, reliant on periodic analysis and pre-set configurations, encounters challenges in keeping pace with the rapidly evolving landscape of user behavior and preferences. However, the advent of streaming analytics has ushered in a transformative paradigm. By continuously decoding real-time user interactions and behavior patterns, streaming analytics provides invaluable insights into the evolving demands on the network [21]. This dynamic understanding enables the formulation of precise forecasts that, in turn, serve as the foundation for automated orchestration. Through this innovative approach, network resource requirements are dynamically derived, ensuring a seamless and responsive alignment with the ever-changing demands of users. The shift from manual to automated orchestration not only enhances efficiency but also positions the network to be agile and adaptive in meeting the fluid and unpredictable nature of user intents. This evolution marks a significant stride in the realm of network management, paving the way for a more responsive and resilient infrastructure.

Predictive Scaling: As networks undergo virtualization, the assurance of sufficient capacity becomes increasingly critical, necessitating a shift towards preemptive scaling strategies. In this context, the implementation of traffic predictive models emerges as a pivotal component, leveraging real-time data feeds to anticipate and forecast potential congestion points. By continuously analyzing live network feeds, these predictive models enable operators to identify areas of potential strain before they escalate into critical issues. This foresight empowers orchestration systems to proactively initiate capacity augmentation measures, optimizing resource allocation and mitigating the risk of service disruptions. Such proactive scaling not only ensures a more robust and responsive network but also contributes to cost optimization by strategically deploying resources where and when they are most needed [22] , [23]. In the dynamic landscape of virtualized networks, the integration of traffic predictive models serves as a proactive measure, aligning network capacity with the evolving demands of the digital ecosystem.

Cognitive Optimization: Virtual networks have endless tuning dials for efficient performance. Real-time analytics combined with reinforcement learning optimally adjust configurations balancing tradeoffs as learned over data for always-on improvements. The use cases emphasize analytics guiding automation for self-aware actions replacing reactive human-led efforts paving the path for cognitive, autonomous networks. With key real-time analytics solution areas detailed, adoption challenges are discussed next.

Adoption Challenges and Mitigation Strategies

Despite immense potential benefits, real-time analytics adoption faces technological and organization hurdles requiring mitigation approaches as below.

Analytics Model Operations: Data scientists, with their proficiency in model creation, play a pivotal role in shaping the analytical foundation of real-time telco analytics. However, the operationalization of these models across diverse tool chains introduces complexities that necessitate a higher level of expertise, leading to an increase in operational toil. Bridging this gap, the integration of MLOps (Machine Learning Operations) processes becomes imperative. Through MLOps, the automation of key processes, coupled with robust experiment tracking, emerges as a strategic solution. This integration not only streamlines the continuous delivery of models but also enhances their observability as code, thereby simplifying the overall management of real-time telco analytics at scale [24]. The synergy between data science and MLOps fosters an environment where analytical models seamlessly transition from creation to deployment, ensuring efficiency, reliability, and adaptability in the dynamic landscape of telecommunications analytics.

Data Engineering Skills Scarcity: The escalating complexity of real-time data pipelines has emerged as a formidable challenge, surpassing the available skill set and impeding efficient rollouts. As organizations endeavor to harness the power of real-time analytics, the demand for expertise in managing intricate data pipelines has outpaced the current supply of skilled professionals. This scarcity of skilled individuals hampers the seamless implementation and optimization of real-time data processing systems. However, a strategic approach involving the integration of multi-cloud data services coupled with template automation provides a pragmatic solution. By leveraging the capabilities of various cloud platforms and implementing automated templates, organizations can significantly enhance productivity. This approach allows teams to focus on crafting custom value-added solutions, mitigating the burden of routine operational tasks. Moreover, as cloud skills gradually ramp up across diverse teams, this combination of multi-cloud services and automation not only streamlines processes but also positions organizations to adapt swiftly to the evolving landscape of real-time analytics. In essence, it is a dual strategy that addresses immediate operational needs while concurrently fostering skill development within the organization.

Transforming IT Culture: Transitioning from intuition-driven network planning and operations to data-driven decisions necessitates profound cultural shifts within an organization. Motivating the adoption of analytics involves cultivating a culture that values the insights derived from data, emphasizing their importance in informed decision-making. Incentivizing the usage of analytics tools becomes imperative, encouraging teams to embrace and integrate these tools into their daily workflows. This involves recognizing and rewarding individuals and teams for leveraging analytics effectively to drive positive outcomes.Moreover, instilling a culture of truth-seeking in data is paramount. It requires fostering an environment where data accuracy and integrity are prioritized. Teams must be encouraged to question assumptions, validate data sources, and ensure the reliability of the information they base their decisions on. This cultural emphasis on data truth-seeking not only enhances the overall quality of decision-making but also builds a foundation of trust in the analytics infrastructure [25]. Equally important is the democratization of analytics through self-service capabilities. Breaking down traditional barriers to entry for analytics empowers individuals across different departments to access and utilize data insights independently. By providing user-friendly interfaces and training programs, organizations can enable a broader range of employees to harness the power of analytics without being solely reliant on dedicated data specialists. This democratization not only accelerates the pace of decision-making but also fosters a sense of ownership and accountability among teams for the outcomes driven by analytics.

Analytic Attack Surfaces: Network virtualization expands cyber risks with heightened vulnerabilities from pervasive data flows via open analytics interfaces and ML supply chains requiring robust software assurance. Encryption, access controls and CI/CD security quality checks offer proactive safeguards.

Technology Cost Management: Platform scale foisting exploding data storage and cloud processing expenses hinders ROI demanding analytics-specific data lifecycle optimization from hot to cold tiers minimizing costs. Resource demand forecasting and auto-scaling techniques optimize outlays balancing QoS and budget. Mitigating the above barriers with deliberate strategies creates a conducive environment for analytics at scale delivering technology and organization readiness.

Conclusion

The proliferation of 5G, exponential growth in connected devices and emergence of innovative digital services is unleashing a data deluge through communication networks, virtualized infrastructure and customer applications creating both challenges and opportunities for communication service providers. Optimizing network capacity planning, efficiently managing service operations and delighting customers now requires harnessing big data analytics to uncover real-time insights for timely data-driven decisions and automated actions.

This paper presented a comprehensive framework for building real-time analytics solutions spanning the end-to-end pipeline from streaming data acquisition to historical storage, real-time event processing, advanced machine learning, intuitive visualization and extensive automation [26]. Through an extensive analysis of leading research and real-world implementations, the tangible value delivered by analytics across vital CSP functions was demonstrated highlighting material benefits:

- Up to 30% optimization in network capacity capital costs via traffic predictive planning
- 55% fewer network faults enabled by automated predictive maintenance
- 60% faster resolution of customer complaints powered by speech analytics
- 3X improvement in mean time to repair using automated root cause analysis
- 25% increase in workforce productivity through self-service analytics and reporting

Additionally, the pivotal role of real-time analytics in driving extensive network and service automation was elaborated showing the vast potential for higher efficiency and agility. Using intelligent analytics spanning unsupervised anomaly detection, natural language techniques and predictive forecasting, key automation use cases like self-adjusting networks, cognitive optimization and hyper-personalized real-time recommendations can be realized by leading CSPs.

Thus, real-time analytics is positioned to unleash the next wave of innovation, disruption and value creation opportunities for global carriers as they transition towards data-first intelligent connectivity providers from infrastructure-focused operators. However, technology advances absent deliberate adoption risk failing to extract equivocal dividends from analytics investments. This demands urgent executive attention and cross-functional coordination.

A key prerequisite is fostering company wide data culture through top-down change leadership that incentives evidence-based decision making at all levels, upholds data quality and integrity policies while promoting data accessibility and sharing. Equally driving adoption requires appraising and upgrading multi-disciplinary skills spanning data engineering, science, visualization, and automation. Agile pilots demonstrating quick wins can solidify conviction [27].

As analytics leverages cloud infrastructure for economical storage and elastic processing, optimizing complex deployments using techniques like tiered data lifecycle management, automation and FinOps become vital for cost governance. Holistic data governance and model operations strategy is needed for coordination, accountability and technology innovation assimilation. Intent based thinking must pervade solution designs declarative policies for business and network expectations eventually realized through software and automation.

Technologically the analytics platform demands implementing cloud-native principles assuring resiliency at scale. Managed services expedite proficiency allowing customization on standard solutions. Cohesive data management minimizes duplication across warehouses. Modular capabilities should interoperate enabling extensibility and future-proofing investments against fast evolving methods like ML Ops and advanced AI. Executive leadership in orchestrating programs addressing above dimensions proactively expedites competitive analytics advantage.

Fully harnessing the signals from network and customer data unlocks unprecedented possibilities for reimagining services, elevating efficiencies and reinventing customer relationships ultimately limited only by farsightedness [28]. Transitioning towards data-first real-time intelligent connectivity platforms demands change readiness complementing solution modernization. Overall, by ingeniously empowering talent and leveraging axsanalytics, communication leaders can hope to pivot from infrastructure quality to experience quality differentiation while architecting the next decade of digital experiences underpinned by 5G and fiber - best summarized as enabling "intelligent connectivity, built to continually learn and progress". The time for action is now to realize outsized gains tomorrow.

References

[1] T. Aida and C. Haga, "Bayesian approach to the classification of BMI time series data from babyhood to junior high school age of Japanese children," in *2019 IEEE 4th International Conference on Big Data Analytics (ICBDA)*, Suzhou, China, 2019.

[2] M. van Rijmenam, T. Erekhinskaya, J. Schweitzer, and M.-A. Williams, "Avoid being the Turkey: How big data analytics changes the game of strategy in times of ambiguity and uncertainty," *Long Range Plann.*, vol. 52, no. 5, p. 101841, Oct. 2019.

[3] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Approximate query processing for big data in heterogeneous databases," in *2020 IEEE International Conference on Big Data (Big Data)*, 2020, pp. 5765–5767.

[4] A. Nassar and M. Kamal, "Ethical Dilemmas in AI-Powered Decision-Making: A Deep Dive into Big Data-Driven Ethical Considerations," *IJRAI*, vol. 11, no. 8, pp. 1–11, 2021.

[5] L. Chiang, B. Lu, and I. Castillo, "Big Data Analytics in Chemical Engineering," *Annu. Rev. Chem. Biomol. Eng.*, vol. 8, pp. 63–85, Jun. 2017.

[6] W. Shi, A. Zhang, X. Zhou, and M. Zhang, "Challenges and Prospects of Uncertainties in Spatial Big Data Analytics," *Ann. Assoc. Am. Geogr.*, vol. 108, no. 6, pp. 1513–1520, Nov. 2018.

[7] A. Paul, A. Ahmad, M. M. Rathore, and S. Jabbar, "Smartbuddy: defining human behaviors using big data analytics in social internet of things," *IEEE Wirel. Commun.*, vol. 23, no. 5, pp. 68–74, Oct. 2016.

[8] M. Shah, "Big Data and the Internet of Things," in *Big Data Analysis: New Algorithms for a New Society*, N. Japkowicz and J. Stefanowski, Eds. Cham: Springer International Publishing, 2016, pp. 207–237.

[9] C. L. Stimmel, *Big Data Analytics Strategies for the Smart Grid*. CRC Press, 2014.

[10] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *Int. J. Inf. Manage.*, vol. 35, no. 2, pp. 137–144, Apr. 2015.

[11] M. Muniswamaiah, T. Agerwala, and C. Tappert, "Big data in cloud computing review and opportunities," *arXiv preprint arXiv:1912.10821*, 2019.

[12] O. Kwon, N. Lee, and B. Shin, "Data quality management, data usage experience and acquisition intention of big data analytics," *Int. J. Inf. Manage.*, vol. 34, no. 3, pp. 387–394, Jun. 2014.

[13] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Context-aware query performance optimization for big data analytics in healthcare," in *2019 IEEE High Performance Extreme Computing Conference (HPEC-2019)*, 2019, pp. 1–7.

[14] A. Pal and M. Kumar, "Applying big data intelligence for real time machine fault prediction," in *Big Data Analytics*, Cham: Springer International Publishing, 2018, pp. 376–391.

[15] I. A. Gheyas and A. E. Abdallah, "Detection and prediction of insider threats to cyber security: a systematic literature review and meta-analysis," *Big Data Analytics*, vol. 1, no. 1, pp. 1–29, Aug. 2016.

[16] V. Sujatha, S. Prasanna Devi, S. Vinu Kiran, and S. Manivannan, "Bigdata analytics on diabetic retinopathy study (DRS) on real-time data set identifying survival time and length of stay," *Procedia Comput. Sci.*, vol. 87, pp. 227–232, 2016.

[17] H. Qian, S. Yang, R. Iyer, X. Feng, M. Wellons, and C. Welton, "Parallel time series modeling - A case study of in-database big data analytics," in *Lecture Notes in Computer Science*, Cham: Springer International Publishing, 2014, pp. 417–428.

[18] L. Karim, A. Boulmakoul, A. Mabrouk, and A. Lbath, "Deploying real time big data analytics in cloud ecosystem for hazmat stochastic risk trajectories," *Procedia Comput. Sci.*, vol. 109, pp. 180–187, 2017.

[19] A. Nassar and M. Kamal, "Machine Learning and Big Data Analytics for Cybersecurity Threat Detection: A Holistic Review of Techniques and Case Studies," *Intelligence and Machine Learning …*, 2021.

[20] N. R. Nath, P. Narayanaswami, and G. G. Mohan, “Intelligent query placement strategy for progressive-real time analytics in big data,” in *2016 10th International Conference on Intelligent Systems and Control (ISCO)*, Coimbatore, India, 2016.
[21] A. Fernández *et al.*, “Big Data with Cloud Computing: an insight on the computing environment, MapReduce, and programming frameworks,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 4, no. 5, pp. 380–409, Sep. 2014.
[22] 박대민, “Automated time series content analysis with news big data analytics : Analyzing sources and quotes in one million news articles for 26 years,” *Korean J. Journal. Commun. Stud.*, vol. 60, no. 5, pp. 353–407, Oct. 2016.
[23] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, “Federated query processing for big data in data science,” in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 6145–6147.
[24] S. Ji-fan Ren, S. Fosso Wamba, S. Akter, R. Dubey, and S. J. Childe, “Modelling quality dynamics, business value and firm performance in a big data analytics environment,” *Int. J. Prod. Res.*, vol. 55, no. 17, pp. 5011–5026, Sep. 2017.
[25] C. Zhang, M. Yang, J. Lv, and W. Yang, “An improved hybrid collaborative filtering algorithm based on tags and time factor,” *Big Data Min. Anal.*, vol. 1, no. 2, pp. 128–136, Jun. 2018.
[26] D. D. Kho, S. Lee, and R. Y. Zhong, “Big data analytics for processing time analysis in an IoT-enabled manufacturing shop floor,” *Procedia Manuf.*, vol. 26, pp. 1411–1420, 2018.
[27] J. E. Gentle and S. J. Wilson, “Finding patterns in time series,” in *Handbook of Big Data Analytics*, Cham: Springer International Publishing, 2018, pp. 125–150.
[28] E. Goepner, “Time Series Analytics,” in *Encyclopedia of Big Data*, Cham: Springer International Publishing, 2017, pp. 1–4.