Challenges in 5G: How to Empower SON with Big Data for Enabling 5G

Ali Imran, Ahmed Zoha, and Adnan Abu-Dayya

Abstract

While an al dente character of 5G is yet to emerge, network densification, miscellany of node types, split of control and data plane, network virtualization, heavy and localized cache, infrastructure sharing, concurrent operation at multiple frequency bands, simultaneous use of different medium access control and physical layers, and flexible spectrum allocations can be envisioned as some of the potential ingredients of 5G. It is not difficult to prognosticate that with such a conglomeration of technologies, the complexity of operation and OPEX can become the biggest challenge in 5G. To cope with similar challenges in the context of 3G and 4G networks, recently, self-organizing networks, or SONs, have been researched extensively. However, the ambitious quality of experience requirements and emerging multifarious vision of 5G, and the associated scale of complexity and cost, demand a significantly different, if not totally new, approach toward SONs in order to make 5G technically as well as financially feasible. In this article we first identify what challenges hinder the current self-optimizing networking paradigm from meeting the requirements of 5G. We then propose a comprehensive framework for empowering SONs with big data to address the requirements of 5G. Under this framework we first characterize big data in the context of future mobile networks, identifying its sources and future utilities. We then explicate the specific machine learning and data analytics tools that can be exploited to transform *big data* into the *right data* that provides a readily useable knowledge base to create end-to-end intelligence of the network. We then explain how a SON engine can build on the dynamic models extractable from the right data. The resultant dynamicity of a big data empowered SON (BSON) makes it more agile and can essentially transform the SON from being a reactive to proactive paradigm and hence act as a key enabler for 5G's extremely low latency requirements. Finally, we demonstrate the key concepts of our proposed BSON framework through a case study of a problem that the classic 3G/4G SON fails to solve.

The advent of the first generation (1G) wireless telephony changed the world by connecting people to people, as its predecessor technology could only connect places to places. Now 5G aims to change the world by connecting anything to anything. Moreover, unlike its predecessors, 5G needs to be conceived as a set of technologies that are efficient and economical in terms of an array of key performance indicators (KPIs) that are of interest to all stakeholders in an omnium-gatherum of applications. These KPIs, from an operator's perspective, include capacity, quality of service (QoS), capital expenditure (CAPEX), and operational expenditure (OPEX). From a user's perspective, the KPIs include seamless connectivity , spatio-temporal uniformity of service, perception of almost *infinite capacity or zero latency*, and, last but not least, the cost of service. Obviously, no technology can offer infinite capacity or zero latency, but by maintaining a latency shorter than the human sensory and physiological delay in the type of application under use, a false perception of infinite capacity or zero latency can be provided. For, example if the network can provide a latency below 100 ms, 10 ms, and 1 ms for audio, video, and tactile applications, respectively, limited by the intrinsic latency of the pertinent human sensory organs and associated neural circuitry, the user will have a perception of infinite capacity and zero latency [1]. However, designing the complete 5G network, only for extremely low latency requirements might it be inefficient in addition to difficult, if not impossible. The more logical approach is to design 5G to be fully *self-organizing* with *end-to-end network behavior intelligence*, from the perspective of a self-organizing network (SON) engine, so that it can exploit the cognition of the context of application as well as that of the state of the network to divert and focus the right amount of network resources when and where needed such that users will perceive seamless and *limitless connectivity*.

5G also has to take into account the recent marriage between Moore's law backed computing power and the wireless technology that has triggered a new era in human history. In this new era the use of wireless communications for novel applications is only bound by imagination. There is hardly an aspect of human life that will not benefit from high-speed

Ali Imran is with the University of Oklahoma.

Ahmed Zoha and Adnan Abu-Dayya are with the Qatar Mobility Innovations Center.

 0890-8044/14/$25.00 © 2014 IEEE

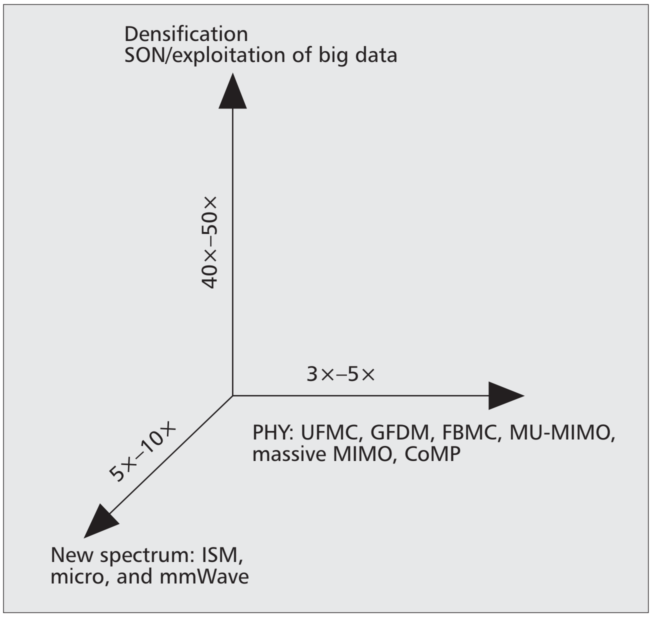

Figure 1. Dimensions and quantification of projected capacity growth in 5G.

wireless communication, including health care, mobility, education, governance, manufacturing, smart grids, entertainment, sports, and much more. The "limitless" and "seamless" connectivity from anything to anything, from anywhere to anywhere, can open new horizons for unforeseen innovations and bring a new level of services and lifestyle to society. Therefore, no single application should be targeted while designing 5G.

Figure 1 illustrates how innovations at the crossroads of the three key thrusts in the emerging wireless landscape can be harnessed together to achieve 600× to 2500× capacity increase for 5G. For a detailed rationale behind these projected capacity growth dimensions, the reader is referred to [2–8]. While these projections promise that capacity targets of 5G are theoretically achievable, it is alarming to note that the complexity of operation, OPEX, CAPEX, and shrinking profit margins are already major challenges for operators of state-of-the-art cellular networks [9]. The typical 2G node has 500 parameters to be configured and optimized, a 3G node has 1000, and a 4G node has 1500. If this trend continues, a typical 5G node is expected to have 2000 parameters.

With this backdrop in mind, analyzing the three dimensions of capacity growth (Fig. 1), we can assume that the operational complexity of the 5G network will scale linearly only with the densification (i.e., the number) of network nodes. The rationale behind this assumption can be that the other two capacity growth dimensions are expected to mostly affect the complexity of user equipment and hardware design. Thus, they may not have a drastic impact on the operational complexity of the 5G networks. Even with these rather optimistic assumptions, to avail of the 40×–50× capacity gain from densification only, if no disruptive measures are taken, we might have to brace for $(2000/1500 \times (40X \text{ to } 50X)) \approx 53$ to 67 times increase in operational complexity and hence OPEX in 5G compared to 4G. The added challenge is that this complexity will be on top of the complexity of the concurrent operation of 2G, 3G, and 4G. Consequently, in contrast to current networks where SON is still a choice and has sporadic penetration so far, for 5G, full embodiment of SON right from the conception level is inevitable to ensure a profitable business model. Furthermore, the fact that major capacity gain in 5G has to come from mostly impromptu densification [8] and network-level efficiency enhancement [8] the technical viability of 5G almost exclusively also hinges on the self-organizing capability of the 5G network. Therefore, the SON paradigm has to evolve radically to enable 5G.

In the next section we highlight the challenges in SON that have to be addressed before it can become capable of enabling 5G. In the subsequent section, we then present our proposed framework for big data empowered SON (BSON), which can not only address these challenges, but play a pivotal role in meeting the envisioned requirements of 5G.

Challenges in SON for Enabling 5G

Fueled by the mounting pressure to reduce OPEX and improve efficiency in legacy networks, the SON paradigm aims to replace the classic manual configuration, post deployment optimization, and maintenance in cellular networks with self-configuration, self-optimization, and self-healing functionalities. A detailed review of the state-of-the-art SON functions for legacy cellular networks can be found in [9]. In the following we only highlight the challenges in SON in the context of 5G.

Underutilized Intelligence: 5G SONs Need Massive Intelligence for End-to-End Network Visibility

Figure 2 explains the generic methodology followed in state-of-the-art 2G, 3G, and 4G SON. Current SON solutions generally assume that the spatio-temporal knowledge of a problem that requires SON-based compensation is fully or at least partially available; for example, location of coverage holes, handover ping-pong zones, or congestion spots are assumed to be known by the SON engine. In state-of-the-art networks this knowledge is obtained through either drive test data, logs of customer complaints, or operation and maintenance center (OMC) reports. [9, 10]. However, this approach cannot deliver the stringent resource efficiency and low latency expected of 5G as it cannot be used to construct dynamic models to predict system behavior in live-operation fashion.

Need for Self-Coordination: 5G Needs Conflict-Free Reliable SON

When operating concurrently in the same network, different SON functionalities can have parametric or objective-based conflicts. Such conflicts may undermine the overall gains of SON. Therefore, self-coordination among SON functions has been emphasized by the Third Generation Partnership Project (3GPP) in order to ensure stable network operation; however, so far it remains an under-addressed problem even for 3G and 4G. From a 5G perspective, given the complexity of the envisioned network architecture, the analysis of potential conflicts generated by the numerous autonomous SON functionalities and the design of an appropriate self-coordination framework can be extremely challenging. Therefore, unlike 4G, where a retrospective approach has been taken to embed self-coordination into relatively independently developed SON functionalities, for 5G self-coordination has to be considered at the grassroots level of the SON functions' design. Reference [11] provides a comprehensive identification and taxonomy of potential conflicts in SON that can be a first step toward designing conflict-free SON functions for 5G.

Need for More Transparent SON: For Full Penetration, 5G SON Needs Operators' Trust

From an implementation viewpoint, the concept of a single SON "black box" encapsulating multiple SON functionalities is very appealing, and therefore has been widely adopted by the

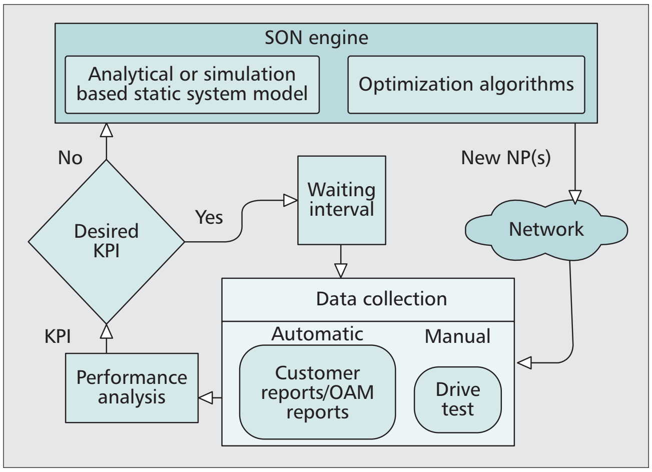

Figure 2. SON in 2G, 3G, and 4G.

vendors for 3G and 4G SON products. However, this approach may threaten the control and transparency network operators want to have in their own network. Therefore, for full penetration, 5G SON has to be designed to offer a certain degree of transparency without compromising the degree of automation. In subsequent sections, we discuss how our proposed BSON framework offers a solution to circumvent this problem.

5G Needs a Focus Shift from Small to Large Timescale SON

Suboptimal performance of wireless networks is mainly caused by the dynamics of their operational environment. The timescale of these dynamics ranges from microseconds (e.g., noise, fast fading), seconds (e.g., slow fading, user mobility), all the way up to hours, days, and months (e.g., change in user concentrations from one area to another over days and nights, and over weekends, events, and seasons). Surprisingly, most of the research effort in wireless communications over the past three decades has been focused on dealing with the inefficiencies and impairments arising from very short-term dynamics (noise, fast fading, shadowing). Arguably, much more system-level efficiency can be harnessed by developing SON solutions for longer timescales. The reader is referred to [9] for further details on operation timescales of SON. Given that not very large gain is expected from the physical (PHY) layer [2], in 5G attention needs to be particularly channeled to design SON to cope with the inefficiencies arising from the lack of network ability to adapt to longer-term dynamics. To this end, a 5G network needs self-organizing capabilities to adapt itself to user concentration changes, user mobility patterns, and data usage pattern drifts over hours of the day, from indoor to outdoor, from weekdays to weekends, and so on.

An Energy Efficiency SON Use Case Requires Exclusive Treatment in 5G

As network densification is the dominant theme of 5G, 5G cells may not be able to afford sticking with the classic *always on* routine. Rather, an operational strategy for most 5G cells might be *turn on when needed*, which is possible only by developing and embedding a whole new set of energy-efficiency-centric SON solutions. The challenges here include defining and designing different levels of energy saving modes for 5G nodes (off, stand-by, hot stand-by, sleep, hibernation, etc.). These operational modes should be able to offer controllable trade-offs to the SON engine in terms of energy efficiency, latency, signaling overheads, agility in returning to an operational stage, and the complexity of self-coordination with other SON functions [11].

The Need for a Holistic Approach: Defining the Right KPIs

So far, the cellular industry does not have a shared framework for network performance quantification. There are as many performance metrics for wireless networks as there are equipment vendors and operators. In 5G, SON is expected to be the de facto operational mode, and SON functions will have to operate on a multivendor, multi-radio access terminal (RAT), and possibly multi-operator and multi-network scale. Therefore, lack of unified performance measurement and a verification system can be detrimental, particularly for 5G SON. The significance of the gain SON can achieve is conditioned on the meaningfulness and uniformity of the KPIs optimized by the SON functions. To realize the full potential of SON functions in 5G, new, holistic, and across-the-board KPIs need to be developed, which can reflect user experience accurately along with quantification of the interests of the network stakeholders.

5G Requires Faster SON: The Need for a Paradigm Shift from Reactive to Proactive SON

SON functions in 3G/4G in general have a reactive line of action; that is, 3G/4G SON functions are designed to kick in when a problem has occurred. For example, a load balancing SON function is triggered when congestion is observed and diagnosed. Given the 5G target of creating perception of zero latency, this type of reactive SON will not be able to meet the performance requirements of 5G. This is because in classic SON, certain time is required to *observe the situation*, *diagnose the problem*, and then *trigger the compensating action*. The resultant intrinsic delay is not compatible with 5G targeted QoE levels. Therefore, for 5G, the SON paradigm needs to be transformed from reactive to proactive. This transformation is possible if, instead of waiting to observe and spot the problem, the problem can be predicted beforehand. This can be done by inferring network-level intelligence from the massive amount of control, signaling, and contextual data that can be harnessed in mobile networks to predict the problem in its infancy, and then take preemptive actions to resolve the problem before it occurs, resulting in a proactive SON. Even if all problems cannot be predicted beforehand, this approach can substantially reduce the intrinsic delay between the observation and compensation phases compared to current state-of-the-art SON. Empowering SON with big data is the key to transforming SON from being reactive to proactive, as we explain in the rest of this articleu.

"Big Data" and Its Utility in Future Networks

What Is Useful "Big Data" 5G SON Can Exploit?

The exact definition of big data is context-specific. In the context of cellular networks, Fig. 3 elucidates and classifies the huge amount of the diverse data that can be available from the mobile network. Given its volume, variety, velocity, and veracity, this data as whole is big data in the context of mobile networks. In the following we further delineate key elements and sources of big data in the mobile network as listed in Fig. 3 by discussing their potential utilities in this specific context.

Identifying Utilities of Big Data in 5G

Subscriber-Level Data — The first column in Fig. 3 lists the data streams, which we label subscriber-level data. It contains control data and contextual data, which not only can be exploited to optimize, configure, and plan network-centric operations, but are equally useful for supporting key business processes such as customer experience and retention enhancement.

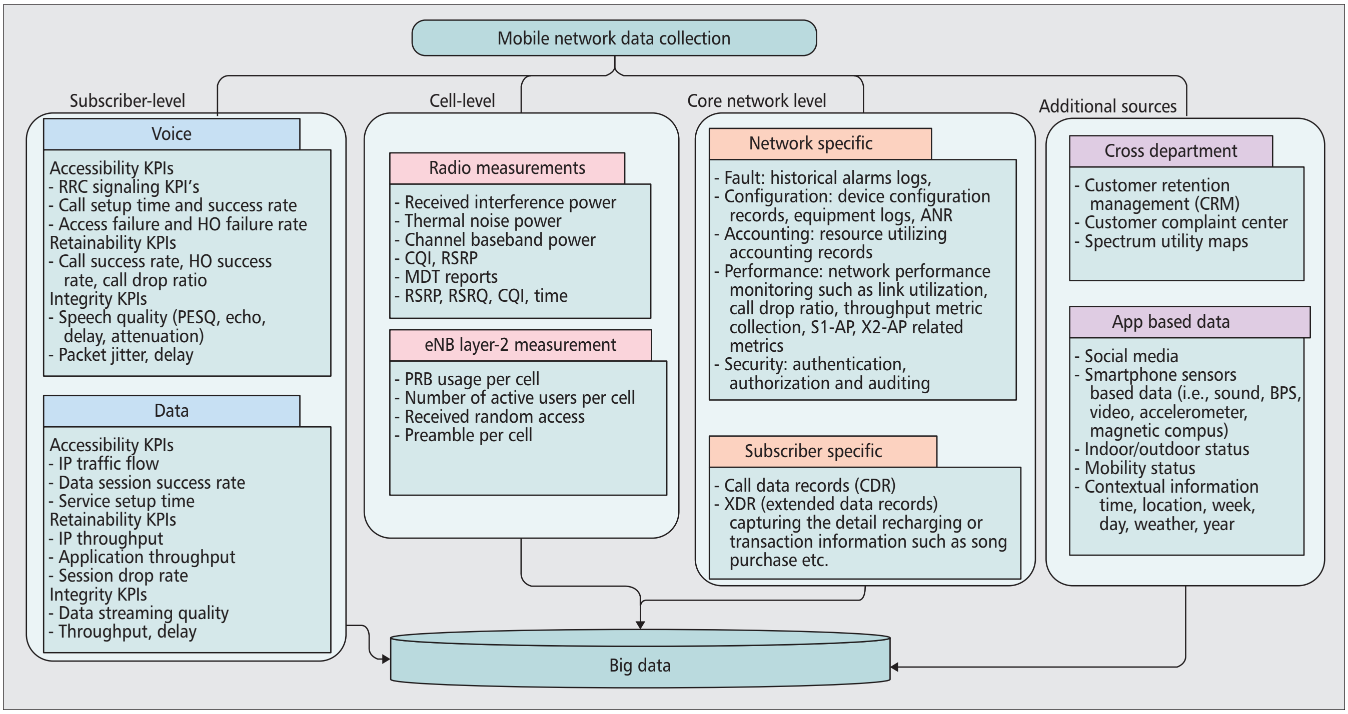

Figure 3. Description of information that can be the part of mobile network big data 5G SON can exploit.

However, conventional approaches relying on offline analysis of individual metrics from these data streams are unable to accurately estimate user experience and create a holistic picture of the network. An up-to-date picture of a network's behavior with fine temporal granularity can only be gathered if information from the subscriber, cell, and core network levels is combined with additional sources of potentially useful data.

Cell-Level Data — The second column in Fig. 3 lists the set of PHY layer measurements that are reported by a base station and all user equipment (UE) within the coverage of that base station to the OMC. We categorize these data streams within big data as cell-level information. The utilities of cell-level data can complement the subscriber level data. For example, Mminimization of drive test (MDT) measurements, which contain the reference signal received power (RSRP) and reference signal received quality (RSRQ) values of the serving and neighboring cells, are particularly useful for autonomous coverage estimation and optimization. The same measurements can also be exploited to develop automated fault detection and localization solutions for identifying coverage holes, sleeping cells, or cells in outage. Similarly, on an aggregated level, layer 2 measurements such as average cell load together with subscriber level data and contextual data such as time, day, and weather information can act as input to the SON engine for performing load balancing, traffic steering, and prediction operations.

Core-Network-Level Data — The third column in Fig. 3 identifies data available from the core network, which is classified as core network data. Data from these streams can be used to fully automate the fault detection and troubleshoot network-level problems. Today such information is only dealt with independently and therefore fails to capture the overall network behavior. The complexity of identifying problems in a core network is increased manifold, particularly if the equipment used is supplied by different vendors that provide their own proprietary solutions to represent network performance. Such problems again advocate the need to share the data from all potential sources at various levels of network operation into a single database, big data.

Additional Sources of Data — Column 4 of Fig. 3 lists the other key elements of big data, such as the structured information already stored in the separate databases including customer relationship management (CRM) as well as billing data. This also includes unstructured information such as social media feeds, specific application usage patterns,and data from smart phone built-in sensors and applications.

The data identified above together can be exploited to gain much more valuable insights into user and system behavior, and develop models that can transform the way a network is operated. Unification of data from multiple sources can also help build models that will be far more accurate than conventional approaches that make use of limited sets of information. For example, management analysis, subscriber behavior analysis, security, and automated troubleshooting are a few potential applications that require information from more than one source. For instance, call detail records (CDRs) as well as extended CDRs can provide detailed information including call duration, and set of statistics at service, bearer and IP Multimedia Subsystem (IMS) level. Together with subscriber- and cell-level information, the CDR can be used to model user mobility behavior. With appropriate privacy preservation measures, the availability of such information within the mobile networks offers the possibility to predict load behavior, discover users' travel patterns, anticipate the whereabouts of mobile users, and further integrate the knowledge with preemptive traffic steering, proactive load balancing, dynamic radio resource and energy efficiency algorithms, and intelligent caching [3], as envisioned for 5G.

A Framework for Big Data Empowered SON for 5G

Figure 4 illustrates our proposed framework to implement big data empowered SON (BSON) for 5G. The core idea of BSON is to develop end-to-end visibility of the network by extracting intelligence from big data through application of appropriate machine learning tools. The three main features that make BSON distinct from state-of-the-art SON are:

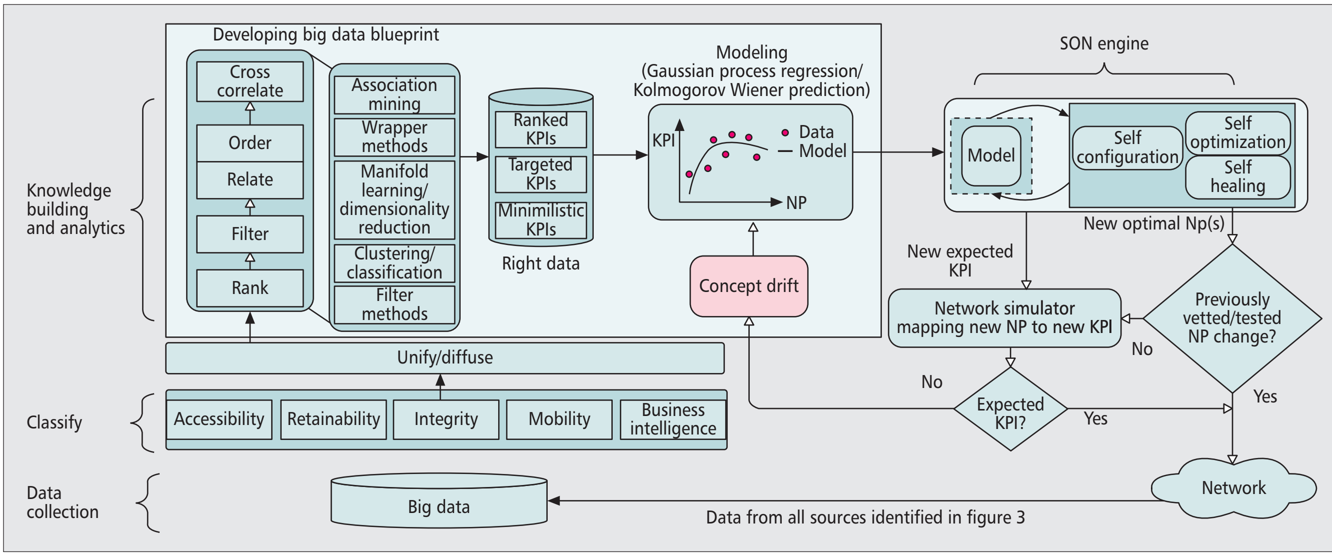

Figure 4. Our proposed BSON framework.

- Full intelligence of the current network status
- Capability of predicting user behavior
- Capability of dynamically associating the network response to the network parameters (NPs)

These three capabilities can go a long way to design SON that can meet 5G requirements. In the following, we explain the operation and functional blocks of the BSON framework.

The framework involves the following steps:

1. **Gather data** from all sources of information into an aggregate data set, big data (Fig. 3).
2. **Transform** big data into the right data by developing its *blueprint*. The knowledge building steps in this transformation are explained below. The underlying machine learning and data analytics are explained subsequently.
   a. **Classify:** Classify data with respect to key *operational and business objectives* (OBOs).
   b. **Unify/diffuse:** Unify multiple performance indicators (PIs) into more significant KPIs.
   c. **Rank:** Rank KPIs within each OBO with respect to their impact on that OBO.
   d. **Filter:** Filter out KPIs that impact the OBO below a predefined threshold.
   e. **Relate:** For each KPI, find the NP that causes an effect on that KPI.
   f. **Order:** For each KPI, order the associated NP with respect to the strength of their association.
   g. **Cross-correlate:** For each NP, determine a vector quantifying its association with each KPI.
3. **Model:** Develop a network behavior *model* by *learning* from the right data obtained from step 2.
4. **Run SON engine:** Use the SON engine on the *model*, to determine a new NP and expected new KPIs.
5. **Validate:** If a new NP can be vetted by the expert knowledge or prior experience of the operator, proceed with changes. Otherwise, determine the simulated behavior of network for new NPs. If simulated behavior tallies the expected behavior (KPIs), proceed with new NPs. (Note that this stage adds the much needed transparency to the black box of the SON as highlighted in section 2)
6. **Relearn/improve:** If validation in step 5 fails, feedback to the *concept drift* block, which in turn will update the behavior model. Even if validation returns a positive outcome, the concept drift block can be triggered periodically to maintain accuracy of the model.

In the following we further explain the steps outlined above.

Developing the Big Data Blueprint: Knowledge Building Perspective

The first step in the knowledge building phase would be to classify the data silos with respect to the OBOs — accessibility, retainability, and integrity of service — corresponding to both QoS as well as QoE, mobility that reflects all data and PIs related to mobility management, and business intelligence, as identified in Fig. 4. The challenge here is not only the immense volume of data but also the fact that it will arrive at different speeds with non-homogenous structure, and may contain incomplete and ambiguous information obtained from different sources. These characteristics of mobile network big data render conventional approaches used for data storage and processing unsuitable. However, there are commercial products available in the market today tailored for telecom operators [12] that offer scalable data processing and management solutions, and thus can be exploited to complete this task for BSON.

Once the data has been classified, the numerous PIs in the data within same category have to be diffused and unified to be reflected as a few selected KPIs for that OBO. A discerning criterion for diffusion and unification is that the KPI formed by the diffusion of the PI is easy to work with while being able to reflect the actual user experience accurately and system performance affectively.

The subsequent step is to perform KPI ranking, which will be the function of the KPI's impact on the OBO category, the significance of that OBO on its operator's policy, and its association with user experience and overall system performance.

Once the ranking is established, filtering operations can be performed to discard KPIs that rank below a threshold. The objective is to minimize the data complexity as well as improve the effectiveness of the subsequent data analytics and machine learning algorithms. The next step is estimating the degree of association of KPIs with corresponding NP, which is a key step toward transforming big data into right data. This is followed by establishing the relationships between different KPIs by ranking their common NP. Since this step enables the system to establish cross-correlation between different KPIs within and across the OBO categories, this not only can assist in self-coordination among different SON functions, but it can offer capability for multi-objective and holistic optimization, with tunable control and transparency. The end-to-end intelligence of the network is another advantage that holistic cross-correlation among cross-OBO KPIs will yield. These features are crucial for smooth SON operation in 5G and are missing from current SON, as pointed out earlier.

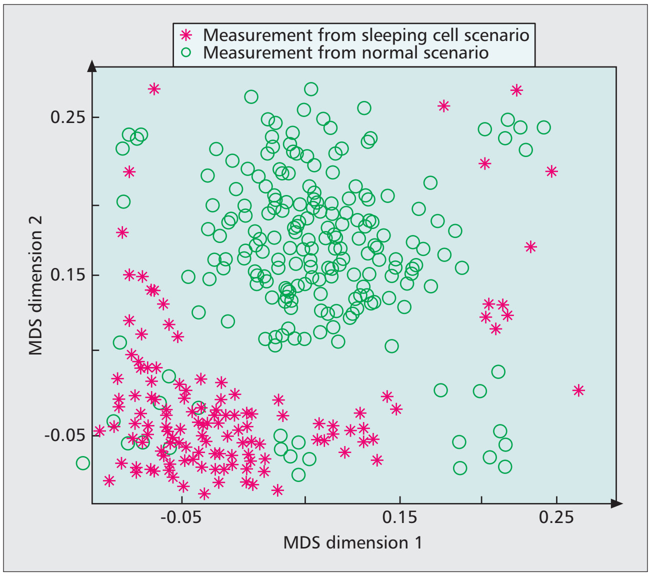

Figure 5. Multidimensional scaling of user radio measurement reports.

Once the cross-correlations between KPIs and NPs have been established, the next step is to develop system behavior models to relate KPIs with respective NPs. In the current SON engine, operators rely on either expert knowledge, offline planning tools, or simplistic analytical models to relate KPIs to NPs. This is not only time consuming and possibly suboptimal given the complexity of the system; it also involves high latency, making this approach unsuitable for 5G requirements. Instead, the BSON framework leverages the right data obtained through steps 2a–g to develop an online system model to both predict and optimize system behavior. This is another key discerning feature of BSON that can reduce the inherent latency of legacy SON and transform it from being reactive to proactive.

Developing the Big Data Blueprint: The Machine Learning and Data Analytics Perspective

In legacy cellular systems expert knowledge is used to analyze the selective network data in individual departments. Given the volume, velocity, variety, and veracity of big data in 5G as envisioned in the BSON framework, manual processing will not be feasible. The "development of blueprint" block in Fig. 4 lists the task-specific machine learning and data analytics tools that are capable of automating the transformation from big data to right data. For example, the filtering process can make use of statistical filtering methods such as Fisher score, Pearson correlation, and information gain to achieve ranked KPIs [13]. On the other hand, wrapper methods can be used to select targeted KPIs by performing KPI subset selection to satisfy the performance criteria along with their network parameters against desired target objectives [13].

To measure the degree of association between KPIs and NPs within and across OBOs, association mining, dimensionality, and manifold learning mechanisms can be applied. In particular, techniques such as principal component analysis and multidimensional scaling can be used not only for filtering operations, but also to establish new attributes by exploiting the interrelationship of high-dimensional KPI sets and further projecting them into a low-dimensional manifold to create a minimalistic representation that can approximate the underlying data well [14]. The key advantage of low-dimension embedding technique is that computational load and latency of the decision making process can be reduced drastically, which in turn can reduce the latency of SON, thus acting as key enabler for proactive SON. (Later we demonstrate the effectiveness of this approach through a case study.)

Modeling

Once the *right data* is there, the next step is modeling of the data to extract system and user behavior models on which a SON engine can rely to perform a variety of SON functions. The classic static methods such as Gaussian process regression, Kolmogorov Wiener prediction, or krigging can be used here, but only for limited SON functions where system behavior does not feature dynamicity. However, since most SON functions actually have to deal with the acute dynamicity of the wireless ecosystem, these classic static methods would not work for those use cases. Below we explain how a recent development in machine learning referred to as concept drift can be exploited to address this problem.

Concept Drift: A Promising Approach for Low Agility, Online, Pro-Active BSON for 5G

The aforementioned problem can be solved with the use of time-evolving stream classifiers that employ a drift detection mechanism using sliding window approaches which process a limited amount of incoming data to detect changes and correspondingly react in real time. Because of the evolving nature of the information, only a summary of this information is stored each time using sketching, load shedding, and aggregation techniques. The summaries are approximate answers of large datasets used for drift detection; in case of a change, the base classifiers are alarmed for an update. There are different approaches to mining data streams with concept drift including ensemble classifiers and option trees. With BSON, these recent developments in machine learning can play a vital role in embedding proactive capabilities in SON for 5G. For example, with this capability BSON can warn in advance of possible abnormality or performance degradation in a cell that is still in the making before it completely loses its operational state. In the next section we present a case study to briefly demonstrate the potential of this concept under the BSON framework.

A Case Study for BSON-Based Detection of Sleeping Cells

A special case of cell outage, referred to as a sleeping cell (SC) [15], is particularly tricky to detect and remains challenging for legacy SON because in this case a cell goes into outage or may perform poorly without triggering an alarm at an OMC. Consequently, no SON compensation function can be launched unless site visits or drive tests are performed, or complaints are received by affected customers. In this section we take this problem as a case study to demonstrate how BSON can solve this problem.

Our solution exploits radio measurements, which are normally discarded after a short time, but in the BSON paradigm can be stored as part of the big , as explained above. Then we show how appropriate anomaly detection models can be developed through machine learning tools to exploit these measurements to intelligently diagnose a sleeping cell. We have set up a simulation scenario consisting of 27 macro sites with three sectors each. In accordance with Long Term Evolution (LTE) MDT specifications, the UEs are configured to periodically report the cell identification (i.e., cell global identity) and radio measurement data (i.e., RSRP, RSRQ) of the serving and neighboring cells to the base station, in addition to event-based measurements whenever an A3 (i.e., a neighbor cell becomes *offset* better than serving) or A2 event (i.e., a serving cell becomes worse than a *threshold*) occurs. The reporting interval is set to be 240 ms, and the simulation is run for 7 min for the two target scenarios, reference and SC. To simulate an SC scenario, the antenna gain of one of the sites is reduced to –50 dBi compared to 15 dBi in a normal scenario. The UE reported measurements, each consisting of nine PIs, are collected from the target scenarios and further stored in a database corresponding to the accessibility OBO. The measurements from the reference scenario are used to model the "normal" network behavior using a k-*nearest-neighbor*-based anomaly detection model. However, instead of using nine PIs per measurements, we extract a minimalistic KPI representation using multidimensional scaling (MDS), which reduces a nine-dimensional dataset to a d-dimensional configuration, whereas in our case $d = 3$. The underlying idea of MDS is that using the information about distances between t data patterns X of dimension n (i.e., which in our case is 9), it attempts to construct t data points $y_i - y_t$ in d dimensions while preserving the inter-point distances. Essentially, MDS tries to minimize the following equation:

$$\min_{Y} \sum_{i=1}^{t} \sum_{j=1}^{t} (d_{ij}^X - d_{ij}^Y)^2$$

where $d_{ij}^X = \|x_i - x_j\|^2$ and $d_{ij}^y = \|y_i - y_j\|^2$. It has been shown in [14] that transforming the distance matrix $\mathbf{D}$ into a cross product matrix X^TX and further performing an eigen value decomposition analysis gives us the solution $Y = \Delta^{1/2}V$, where Δ and V are top eigenvalues and eigenvectors of X^TX, respectively. The minimalistic KPI representation exploits the interrelationship between PIs to construct an embedded space where similar measurements are placed close to each other, whereas the dissimilar measurements representing an anomalous network behavior are projected far as shown in Fig. 5. This allows the learning model to autonomously and dynamically profile network behavior with high accuracy. Our analysis shows that BSON based detection model has achieved a detection accuracy of 94 percent, when tested against the measurements belonging to the SC scenario, with as short training time as 7 min. This is a great improvement over legacy SON, where SC can go unnoticed for hours and days. This preliminary study demonstrates how similar exploitation of BSON can go a long way to realize the goals of 5G.

Conclusion

While the SON paradigm has evolved over the past decade to automate 2G, 3G, and 4G, we explain why it may not meet the requirements of 5G, mainly, because of its intrinsically reactive design approach and lack of end-to-end knowledge of the network. To address these problems we have laid down a vision for empowering SON with big data. We provide a detailed framework for implementation of big data empowered SON (BSON) in 5G. We detail the deluge of largely untapped data that can be harnessed in future cellular networks to realize BSON. We explain how well established and powerful tools from the domain of machine learning and data analytics can then be leveraged to structure, analyze, and utilize this information to create end-to-end visibility of the network to implement SON that is faster and more transparent, and can be proactive instead of reactive, thus meeting the diverse and acute 5G requirements including extremely low latency. We also demonstrate the viability of proposed ideas through a brief case study. Additionally, the vision laid out in this article has two further ramifications in the context of 5G. First, BSON in its broader manifestation can allow service providers to create new business models by monetizing the knowledge base gained by the non-intrusive user-based profiling for applications in vertical sectors such as health care research, transportation, urban planning, marketing, governance, security, and administration. Second, we imply that to realize the projected benefits, BSON requirements need to be incorporated into 5G design and standardization at its very earliest stage to ensure the availability of sufficient and necessary data without compromising user privacy.

Acknowledgment

This work was made possible by NPRP grant No. 5-1047-2-437 from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors. More information about this project can be found at www.qson.org.

References

[1] G. P. Fettweis, "A 5G Wireless Communications Vision," *Microwave J.*, Dec. 2012, pp. 24–36.
[2] Q. C. Li *et al.*, "5G Network Capacity: Key Elements and Technologies," *IEEE Vehic. Tech. Mag.*, vol. 9, no. 1, Mar. 2014, pp. 71–78.
[3] X. Wang *et al.*, "Cache in the Air: Exploiting Content Caching and Delivery Techniques for 5G Systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 131–39.
[4] T. Alsedairy *et al.*, "Self Organizing Cloud cells: A Resource Efficient Network Densification Strategy," *Trans. Emerging Telecommun. Technologies.*
[5] T. Taleb and A. Ksentini, "Follow Me Cloud: Interworking Federated Clouds & Distributed Mobile Networks," *IEEE Network*, vol. 27, no. 5, Sept./Oct. 2013. pp. 12–19.
[6] G. Wunder *et al.*, "5GNOW: Non-Orthogonal, Asynchronous Waveforms for Future Mobile Applications," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 97–105.
[7] T. Taleb, "Towards Carrier Cloud: Potential, Challenges, & Solutions," *IEEE Wireless Commun.*, vol. 21, no. 3, June 2014, pp. 80–91.
[8] N. Bhushan *et al.*, "Network Densification: The Dominant Theme for Wireless Evolution into 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 82–89.
[9] C. G. Aliu *et al.*, "A Survey of Self Organisation in Future Cellular Networks," *IEEE Commun. Surveys & Tutorials*, vol.15, no.1, 1st qtr. 2013, pp. 336–61.
[10] I. Ali *et al.*, "Self Organization of Tilts in Relay Enhanced Networks: A Distributed Solution," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, Feb. 2014, pp. 764–79.
[11] H. Y. Lateef, A. Imran, and A. Abu-dayya, "A Framework for Classification of Self-Organising Network Conflicts and Coordination Algorithms," *Proc. IEEE PIMRC '13*, 2013, pp. 2913–18.
[12] Amdocs Smart NetSolution, http://www.amdocs.com/Products/OSS/Pages/small-cell-solution.aspx, visited on 17 Apr. 2014.
[13] I. H. Witten and F. Eibe, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, 2005.
[14] T. F. Cox and M. A. Cox, Multidimensional Scaling. CRC Press, 2010.
[15] S. Hämäläinen *et al.*, *LTE Self-Organising Networks (SON): Network Management Automation for Operational Efficiency*, Wiley, 2012.

Biographies

ALI IMRAN is an assistant professor in telecommunications at the University of Oklahoma. He is currently leading a multinational $1.045 million research project on self organizing cellular networks, QSON (www.qson.org). His research interests include self-organizing networks, radio resource management, and big data analytics. He has authored over 30 peer reviewed articles and has presented number of tutorials at international fora such as IEEE ICC and WCNC on these topics. He is an Associate Fellow of the Higher Education Academy, United Kingdom.

AHMED ZOHA completed his Engineering and Physical Sciences Research Council (EPSRC) funded Ph.D. degree in electronic engineering from the University of Surrey, United Kingdom, in 2014. He is currently working as a research scientist at the Qatar Mobility Innovation Center (QMIC), Doha, focusing on self-organizing network solutions, particularly automated fault detection and self-healing aspects of cellular networks. His research interests include machine learning, big data analytics, knowledge discovery, and self-organizing networks.

ADNAN ABU-DAYYA received his Ph.D. in electrical engineering from Queens University, Canada, in 1992. He has more than 25 years of global experience in the areas of ICT R&D, product innovations, business development, and services delivery. He is currently appointed as the founding executive director (CEO) of the Qatar Mobility Innovations Center (QMIC) at the Qatar Science & Technology Park. He also serves on the advisory boards of the Arab Innovation Network, and the Department of Electrical and Computer Engineer at Texas A&M University at Qatar.