PART II

A NEW PERSPECTIVE

"Imagine a vast sheet of paper on which straight Lines, Triangles, Squares, Pentagons, Hexagons, and other figures, instead of remaining fixed in their places, move freely about, on or in the surface, but without the power of rising above or sinking below it, very much like shadows - only hard and with luminous edges - and you will have a pretty correct notion of my country and countrymen. Alas, a few years ago, I should have said "my universe": but now my mind has been opened to higher views of things"

- A. Square in "Flatland, A romance of Many Dimensions" by E.A. Abbott.
Chapter 5

Problem Analysis

A fully automated production street, such as a hot strip mill, is the result of evolution. Connecting industrial machines progressively, the systems global function and qualities emerge raising the need for integrated plant operation. Communication infrastructure and services evolve to ever larger and increasingly linked systems; their global quality aspects are complex functions of the quality of its components and its diverse usage. Some remote sensing and controlling networks, such as LOFAR, are intentionally linked to achieve a global function through coherent signal processing. These systems are very different from conventional systems addressed by Fault Detection and Isolation. New and more adequate organizations of control appear, but still unforeseen disturbances defy attempts to contain and localize abnormalities, resulting in systematic errors and harmful fault propagation. There is an obscure gap between the nature of systematic deviations and the capabilities of classical theoretical detection approaches. A better understanding is required. We need to answer: "What are the challenging disturbances?", "What is the origin of these disturbances?", and "Why are conventional methods inadequate?".

This chapter is a problem analysis, following a method of deriving the problem statement from the phenomena through an iterative refinement of our understanding of the limitations of conventional detection methods. First we clarify the application domain addressed in section 5.1. We discuss the phenomena observed in real-world cases in section 5.2, arriving at a better understanding of how locally autonomous distributed systems and abnormalities differ from the conventional detection problems. Section 5.3 provides the in-depth problem analysis, focusing us on the actual problems.

5.1 Applications in distributed systems

In the section we consider general aspects of Locally Autonomous Distributed Systems (LADS). Autonomous local processing evolves from direct control through automation and the expansion of systems. We first summarize the key properties of controlled systems from the perspective of detection, followed by the key trends in the emergence of distributed sensing and control systems. Sensor networks are described. We conclude with the key aspects of LADS.

A potential for computational intelligence

Industrial production systems manifest disturbances that make the desired production quality and availability hard to reach. Consequently, adequate fault detection and accommodation are required complementary to the already automated control in production streets, such as in extruder pipe machinery and hot-strip mills. These systems manifest a complex behavior which cannot be understood fully even though the utilized physical and logical principles are well understood in isolation; particularly the effects of wear and tear, evolving use and varying
conditions complicate the behavior. They are truly time-varying systems. The deviations, specifically time-related disturbances, from the blueprint may be interpreted as patterns by a human, but in real cases they are of a seemingly erratic roughness not fitting to any simple mathematical or statistical model. Nonetheless, they appear as patterns to the human eye, and the disability of conventional purely theoretical approaches to detect these patterns is all too often an unacceptable status quo.

Despite the physical and mathematical foundation of the disciplines involved in systems design and operation, the pragmatic industrial R&D sections have opened up to less conventional techniques, i.e. computational intelligence, to complement the existing arsenal. Computational intelligence includes quantitative methods such as neural networks, fuzzy logic and evolutionary algorithms. Our research originates from this setting, i.e. to investigate the potential merits of neural networks to detect and accommodate time-related disturbances in batch-oriented processes.

**Grids and sensory networks**

We focus on health monitoring through early detection for grids and sensor networks, such as energy Grids, wide area communication systems, transport and environmental monitoring infrastructures. The latter are of increasing importance considering humanity's footprint on mother earth [IT Roadmap for a Geospatial Future, 2003]. The objective of sensor networks is to guarantee the reliable operation of distributed systems essential to industrial plants or environments. The correct operation of such networks is of increasing importance to our society. Sensor network applications aim to identify and govern physical processes, systems or ecologies through a network of distributed sensors with locally autonomous processing capacity. The reliability of sensor networks is to be guaranteed on different abstract levels. A combination of platform monitoring and signal inspection can provide such means. Governing locally independent processes for global requirements depends on a shared and consistent model. The efficient synthesis and dissemination of this model for validation of local correctness is a major research challenge.

**Applications: global functions and qualities, and evolution in use**

The common divider of many networked applications in industrial automation and sensor networks is the application domain of LADS with a global function and qualities. Though the system becomes modular the functions and qualities of the system are not as partitioned as the modularity of the system may suggest. In sensing networks the sensing as well as the processing and storage is highly distributed (even physically), yet a globally accurate model of the physical phenomena has to be construed. Sensor networks observe processes that are governed by similar laws originating from similar sources or purposefully designed physical behavior. Coherency and similarity in distributed behavior provides means for detection and diagnosis from within.

Industrial processes, sensing networks and communication systems have life-cycles that range over more than a decade. In practice the requirements and conditions are evolving while a system is in its operational phase. Particularly for science instruments, their use is evolving and these systems are pushed to their limits; progressive upgrades are a result of this. The waterfall approach has lost its attraction, as the requirements and functions of the systems cannot be frozen at the design time.
Technology enabled trends

There are three technological trends in the capabilities of systems that have become increasingly distributed and increase in scale. These trends are:

- **Sensors/actuators.** The miniaturization and digitization of sensor and actuators enables decentralization of control: expensive sensors are replaced by multiple cheap ones combined with intelligence for integrating sensor measurements into reliable estimates. Computational power becomes cheaper than the material cost for high-quality sensors. Centralized signal processing is slowly dissolving, when the individual sensors become more and more intelligent.

- **Bandwidth.** Availability of communication infrastructure enables linking of distributed processes. Systems can share diverse information and measurements across large areas.

- **Automation.** The evolution of systems shows a trend of increased connectivity of processes. There seems to be an on-going extension of systems and models to improve performance and quality using new technologies. Systems evolved without blue-prints, particularly existing systems are pushed to achieve better performance and varying use. A consequence of automation in the design and engineering is re-use of both hardware and software components. This results in an inherent similarity in the behavior causing faster replication of weaknesses\(^1\), and it causes a coherency in emergent behavior where it may be undesirable.

The technological advancements are key enablers for increased capabilities. In isolation the technological advancements do not necessarily increase complexity. However, the consequences of technological advancements are decentralization, distribution and automation. These in turn increase the complexity, particularly the organization of the system control.

Organizations: the structure of control

Classical detection approaches discussed in chapter 4 are predominantly centered on control-oriented systems. Understanding the physics in the system inspires design of control models from these physical principles, and is based on the belief that set-points are computable from the state and observation. The pursuit of an equilibrium through control, which is paramount for stability in production quality, results from a view that off-sets and disturbances are, in absence of abnormalities, quasi-stationary time-series. It is not obvious that the principles in classical detection are still suitable when systems become more autonomous. Gradual increments in scale, distribution and automation of systems call for gradual increments in the control organization of the system. We distinguish the following gradations:

- **Direct Feedback control.** Control over mechanical, physical and chemical processes is the application domain of dynamic system theory, relying on physical-principle models. In direct feedback control the sensor data is directly used, via state observers and controllers, to compute control stimuli. Aeronautics is the typical example: fly-by-wire.

- **Hierarchical direct control** is the central direct control over local controllers. An examples is hierarchical PID-control. The transitions of state A to state B at any hierarchical level are directly guided centrally, while the granularity of the global control and local controller only vary slightly. Intermediate control actions are computed according to a

---

1. The number of bugs per 1000 lines of code had remained approximately constant in the past decade.
deterministic procedure, running a batch program and a preprogrammed response to wear. A production street such as the automated hot-strip mill falls in this category. A key criterion of this category is a *central procedure specifying "how to get from A to B"*.

- **Hierarchical set-point control.** The desired state of the system as a whole is related to the desired state of local processes. The key criterion of this category is that the local controllers are *autonomously determining "how to get from A to B"*, and the global transition is translated to local set-points. This introduces local autonomy. The control strategy of this category accepts the impossibility to centrally control the behavior of processes at the lower level at their finest grain. Sensor networks such as LOFAR are in this category.

- **Self-organizing systems** are characterized by the specification of a desired function or results while the internal state of the system is not specified anymore. The results are achieved through self-organization, if it is achieved at all. Control mechanisms for self-organization are partly defined and partly implicit. An mechanism competition over resources (information, communication bandwidth and computing time) which effects interactions in an ecology of processes. The processes, or agents, are autonomous, but have a common "blueprint", a set of rules in their "genes". The communication service network is in this category.

- **Ateleological systems** lack an apparent goal or purpose, as the result is not specified. The behavior of the system as a whole has implications for the "processes" it contains. There is no external regularization; the impact of intentional action is not predictable. Think of ecosystems!

Systems have increasingly become more complex raising the need to abandon direct control. Less coherent system organisation implies a more complicating fault prevention. Particular functions that are implemented on these systems are networked applications: the function arises from a collaboration of many separate nodes rather than from a single entity. The relation between functions and the system platform becomes less clear, subsequently the fault detection and accommodation becomes increasingly complex.

**Key aspect of locally autonomous distributed system**

LADS have multiple distributed processes and a modular/hierarchical control. Key points in the design of the control strategy in LADS, are: it is hierarchical, and the required result (set-point) is specified rather than the behavior required achieving it. Systems are modular and lack a single control-interface, as the system processes are more autonomous. However local autonomous processing does not guarantee correct global behavior of the system. A single process can still be understood from it’s initial design; the behavior of cooperating processes is not trivially explained from the behavior per process. There are global as well as local conditions and controller interactions. The physical and logical principles are incomplete, preventing a clear understanding of the relation between the design of the system and the manifesting behaviour.

**5.2 Inspiring phenomena**

A number of case studies has been performed in the line of research that will be reported in this thesis. The examples actually represent steps in the evolution from industrial processes with distributed control to collaborative systems with highly autonomous local nodes; an evolution where the composition of the function is increasingly complex, and the control increasingly hierarchical. Table 5.1 characterizes the cases according to these aspects. Detection challenges
emerge because these systems are locally autonomous and support one or more distributed applications. These are man-made systems developed for particular purposes at some time, but have technologically advanced beyond manageability by conventional means.

Table 5.1: Classification of cases by typical properties of distributed systems

<table>
<thead>
<tr>
<th>case</th>
<th>type</th>
<th>scale [km]</th>
<th>multiplicity</th>
<th>parallelism</th>
<th>hierarchy</th>
<th>control organization</th>
<th>local controller</th>
<th>central clock</th>
<th>local clock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot-strip Mill</td>
<td>Plant</td>
<td>0.5</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>Hierarchical direct PID control</td>
<td>Real-time controller</td>
<td>s</td>
<td>ms</td>
</tr>
<tr>
<td>Communication Service Network</td>
<td>Network</td>
<td>0.01</td>
<td>100</td>
<td>1000</td>
<td>1</td>
<td>Monitored Self-organizing</td>
<td>local operating system</td>
<td>minutes</td>
<td>ms</td>
</tr>
<tr>
<td>Sensor Network</td>
<td>Sensor Network</td>
<td>150</td>
<td>1000</td>
<td>10000</td>
<td>5</td>
<td>Hierarchical set-point control</td>
<td>Streaming data</td>
<td>ms</td>
<td>us</td>
</tr>
</tbody>
</table>

5.2.1 Industrial plant: a hot strip mill

Application

This is an example of quality and performance improvement in highly automated industrial production, manufacturing and assembly plants. It is the classical production street as popularized by Ford, whereby material goes in at the one side and the product comes out at the other. In a different view, it is a linear arrangement of individually controlled machines. Much of the data and experimental evidence we quote in this thesis comes from an analysis performed for Corus IJmuiden [vanderSteen, 2001].

System

A hot-strip mill has a very descriptive but also somewhat misleading name. A long strip of heated metal is rolled through a series of mills (rather than just one) and thereby successively pressed to the desired thickness, simultaneously making it longer. In the process of pressing the steel plates the plates decrease in thickness from O(10cm) to O(1mm), and consequently the length and width increase, particularly the length increases from O(10m) to O(km). Consequently the initial pace of a few m/s increases to O(100km/hr), and the force control need to respond rapidly to variations. The variations are measured directly after each of about 10 mills. Hence each mill has its own control, set to a value that is derived from the global target by means of a physically plausible model for pressing metal shapes. The hot-strip mill operates on batches: operation is started, comes on steam, produces the lot, slows down and stops. All these different modes and the shady regions in between must be handled by the controller, or rather the global control task has to be divided into subtasks on individual mills in potentially different modes of operation.
Problem: manifesting abnormalities

The hot-strip mill has evolved in a few decades. Particularly in the last decade the market challenge pursued by the steel industry is to minimize thickness variance of the steel strip and diversity in materials, even hybrids. A more general ambition is to reduce the loss of material, i.e. the head and tail of each strip is deformed, looking like a fish tail, due to initialization and termination problems. The increased demands on quality particularly apply to the variance in the thickness. The material loss and strip profile control depends at first on the direct control, which resembles a PID-controller. However the main variable to be optimized is the initial set-point of the local PID-controllers. The manifested disturbances are the off-sets of these initial set-points causing the fish-head as well as variations (controller settling effects). The set-points are estimated or rather predicted by a combination of models. The off-sets or deviations vary over a batch due to wear of the mills, variations in conditions and variations in processed material. Since the PID-like controllers correct the force applied during the processing, the deviation from the chosen target force is measurable. One measurement per stand per plate is thus collected, not sampled uni-distant. The predictions of different models, conditions and configurations are all available in huge amounts of data, taken over a few years. The disturbances are dynamic in nature, and the challenge is to improve the role-force prediction through a corrective model on the existing models.

Modeling the system

The operation of each individual mill is largely understood on physical principles but not in sufficient details for a complete analytical model. Adaptations are necessary because all mills are slightly different in construction and react different to ageing. Statistical measures are popular to construct and maintain such a process model. Not always does the applied statistical technique provide better understanding of the physics. A regular stream of corrective models has been necessary over time to answer the increasing demands in production efficiency and environmental dependability. The modeling concept gets more confused for the overall milling street. The cross-terms are not easier to determine by hand than by a black-box model. A global model is required in addition to local stand models.
Solution approaches

Dynamic variations in the process and time-related disturbances can also be accommodated by applying neural methods. We have investigated the possibility to make instant corrections per stand and for the whole street simultaneously. There are dynamics on very different time-scales, i.e. slow variations due to wear and short-term variations due the material variations and such. We also consider two types of dynamics in the model to capture and accommodate the variations over time: dynamic neural networks and on-line learning. On-line learning neural networks have been applied to accommodate the slowly varying time-related disturbances. The experiments reveal that the processes themselves must not be treated independently; inter-process dependencies need to be considered, while moreover a vast amount of exceptions come into play. An artificial neural network (ANN) model of combined mill data, fitted to the whole pipeline across all instances, significantly improves the set-point prediction. Furthermore a multiplicative or additive correction of an existing physical-principle model is outperformed by a direct prediction from the ANN. Still it has been useful to add the prediction made by the mathematical and statistical models, particularly some statistical cross-terms. A trend analysis of the residuals has revealed that disturbances have dynamic patterns across various time-scales, varying from minutes to hours. Some of these patterns can be accounted for by the maintenance schedules. The on-line learning configuration is a key design issue, as the plasticity and stability of the neural model have to be tuned to the dynamics in the data. A few time-steps and fading memory with low resolution are helpful. However time-series models on the resolution of the sampled data for the trends within and across batches are not useful. Particularly the non-equidistant sample is a complication for time-series modeling and on-line adaptations appear to be more capable of accommodating the temporal variations.

5.2.2 Network services: communication

Application

The integrity, confidentiality and availability of networked communication services is of vital importance. This is an example of network-related problems that bother data communication infrastructures supporting business and governmental transactions. We refer here to analysis performed within and for KPN Research\(^1\) [Hut, 2000]. KPN provides a range of services on fixed and mobile networks. The quality of the services depends much on the quality of the software design in the network service platform. A particular issue is security. It appears to be impossible to track down and eliminate every hole in the system since software bugs are unavoidable, design flaws as well as configurations and interactions whose potentially weak behavior can't be foreseen from the systems design. Consequently monitoring is required.

System

Communication services are mapped onto networks consisting of many interconnection networks, routing subsystems and computing and storage nodes. The subsystems and network services are increasingly decentralized and they are owned by various companies. It is a particular complication that network boundaries become less and less clear. There are many ways to connect and the network platform evolves beyond it's original size and owners.

Problem: Manifesting Abnormalities

---

1. Currently TNO Telecom
Unauthorized access by both insiders as well as outside attackers can do direct harm by taking out computing nodes that are part of the networks platform or simply by pre-emptive use of resources at the cost of the quality of services for authorized use. The abnormality either manifests as a known session trace (in which case it can be diagnosed early) or it goes undetected until the harm is done (then the causes can only be analyzed in retrospect). Expert systems and diagnostic systems for misuse detection are only as good as the security operator whose input is used. The knowledge base is never sufficiently up-to-date to allow for prevention of harm. The problem is to make a distinction between normal and abnormal intrusive network behavior without relying on a priori knowledge. Abnormalities are observed from either (a) the availability of resources in the platform if it is adequately monitored; or from (b) the data packets in the network traffic; and from (c) the logs that are kept on the hosts and routers in the network. Of course in practice the challenging abnormalities to be detected are so far only detected by an out-of-service notification. Apparently stochastic data-driven modeling is called for, despite the deterministic nature of the digital platform and the software components.

Modeling the system
A model of the system can be based on the structural models that are at the basis of the system’s design (infrastructure, protocols stacks, state and interaction diagrams and exception handling). Alternatively models can be acquired from the data collected from the hosts and routers or the network traffic. The volume of the data is a complication here, and data reduction is essential. Though hybrid models from data as well as design models have been pursued, the coupling between the design models and a data-driven model is difficult as soon as the necessary data reduction is used.

Solution approaches
A solution approach is self-learning artificial intelligence for abnormality detection; or rather novelty detection since it aims at distinguishing a priori unknown abnormalities. Another approach is agent based detection. The research of H. Hut [Hut, 2000] focuses on self-learning artificial intelligence, and at the time of this research the detection of intrusions directly from network data was ambitious. The detection is based on a data-model reflecting the normal behavior, where the data model is a clustering of packet features. The clustering is obtained and maintained through a self-organizing feature map. The network traffic has subsequently been inspected by looking at the deviations of the trace from it’s cluster in the feature map, that it is the error of the ordered data-packets without considering the address information. A straightforward inspection follows from the time-ordered error-signals over time, including all unknown simultaneous communication sessions and integrating over all the clusters in the model. Already a graphical inspection of the error-traces through time revealed distinct session traces.

We can detect automatically, by means of error-variance threshold, one out of four types of intrusions. Though we cannot publish the actual data, the figure 5.2 (left) shows a typical graph of the error-signal for blindly collected network traffic. The curvature of the variance in the error-signals offers the possibility to compare normal network traffic with abnormal traces, as illustrated in the figure 5.2 (right). The black-box nature of the model prevents an interpretation of the error data as is; nonetheless the isolation of the session traces is possible with the human eye and can be further pursued. Surprisingly a service providing system of independent machines running deterministic protocols manifests seemingly random behavior; an interesting
discussi omission of this phenomenon is given in Linked [Barabasi, 2001]. The behavior of communication services reveals dynamic patterns on a macroscopic scale. Moreover a dynamic model of the session traces does not require many parameters, and its estimation from data allows for detection of abnormality as well as prediction.

It appears to be possible to distinguish normal from abnormal session traces, allowing more intrusions to be detected. A data-driven model of network behavior appears to be possible and without a priori information. It appeared hard to classify normal behavior from intrusions, and the modeling approach is perceived as being too complicated for practical use. Abnormalities, are to be considered as time-related disturbances for detection. Note that the traffic at one location is still a manifestation of global interactions.

The many protocol layers in systems are to shield issues typical for a certain layer from higher protocol layers. Deterministic models are at the heart of the design. Yet problems occur and the system does not behave deterministic/predictable. Software engineering discipline is keen on formal methods and keeps trying to get to a complete and consistent model of reality to drive the software design, and fail. The KPN case demonstrates that an inherently deterministic system ought to be analyzed in a macroscopic stochastic way. Today’s QoS issues are still addressed from within the system using deterministic rule-based and agent-based software solutions. Dynamical global behavior is hardly used to facilitate early-warning.

5.2.3 Sensory networks: low frequency array

Application

The Low Frequency Array (LOFAR\(^1\)) is an instrument supporting various applications in astronomy, geophysics, and agricultural monitoring. Astronomy addresses the fundamental questions "where does it all come from?" and "what will be the future?". These questions have enticed humanity to investigate our universe; therefore it is the mother of all sciences. Primarily LOFAR aims at breakthrough discovery in radio astronomy at low frequencies (30-240MHz) as compared to the dominant frequency in the universe of 1.4 GHz, the radiating frequency of hydrogen. One application is the discovery of the Epoch of Re-ionization, an era in the history of the universe which according to the big bang theory should exist and would reveal a global phase change in the Universe caused by the appearance of the first luminous

---

1. http://www.lofar.nl
objects. Furthermore LOFAR facilitates the study of transient radio sources generally. Also the solar studies are worth mentioning as the nearest star still has many secrets and the prediction of solar winds can help to protect electronic equipment. It will allow also study of high-energy particles. The traditional areas of astronomical observation are imaging and spectral analysis. LOFAR provides these with a much higher sensitivity, a multi-beaming capability and an instantaneous bandwidth of 32MHz, comparable to monitoring 1000s of radio channels in many respects. LOFAR will open a new window on the universe nearby and very far away.

**Figure 5.3**: a) the LOFAR station antenna field; b) LOFAR station distribution

**System**

ASTRON is developing LOFAR as a distributed sensor array divided into three main sub-systems: the stations, the Wide Array Network and the Central Processor. The sensors are small antennas rather than parabolic dishes. The "dishes" are formed digitally through a process called beam-forming. The processing is derived from array processing with parabolic dishes. These reflector arrays have been introduced in the 1960s. The Westerbork Radio Telescope Array has brought the technique of interferometer imaging and the essential calibration techniques to achieve the sensitivities required for astronomical discovery. The shift from dishes to antennas is a major paradigm shift in radio-astronomy. Particularly since the instrument now becomes omni-directional, i.e. the beams are formed digitally and the number of beams formed depends only on the processing and communication capacity that can be afforded. In the area of astronomical instrumentation that is traditionally purely sensitivity driven, the provided flexibility poses enormously challenges particularly to calibration, but also for the astronomical community to benefit from these capabilities. The system is a highly distributed parallel hierarchical signal processing machine. There are about 100 stations each with 300 digital processing components, a wide array network, and a huge central processing facility with an IBM Bluegene/L and additionally Linux based cluster computers. These platforms support an enormous variety of signal processing and computing processing. On one station there are hundreds of signal paths with various filtering and beamforming stages executed in parallel on FPGAs, controlled hierarchically on 1ms accuracy.

**Problems**

In the context of this thesis there are two aspects of interest in the LOFAR design. The first aspect is stability of the sensitivity of the instrument and therefore its calibratability. To guarantee the required sensitivity for astronomical applications instrument behavior in terms of self-generated noise and man-made radio-frequency interference need to be well understood.
This requires in-depth understanding of the local technology-specific gain and phase variations, e.g. those introduce by the ionosphere, clocks jitter, antenna radiation patterns, low-noise amplifier behavior, A/D stability, multi-rate digital filterbank limitations such as aliasing and rounding; and beam forming limitations such as side-lobes. Note however the network latency variations are no impact on signal quality as the data is time-tagged. All mechanical, analog and digital components need to behave such that the system is sky-noise limited, i.e. the system noise is less than the very weak signal of celestial sources across a wide frequency band with an instantaneous band of 100 MHz. The ionosphere introduces local gain and phase variations whose effect on the global performance of the system are not known a priori. All these effects deviations from the ideal function, understanding through a thorough analysis pursues optimal use of technology and architecture of the instrument. However the impact of design choices is difficult to quantify on a system level, which poses a major threat to the convergence of the system design. It has become apparent that a divide-and-conquer hierarchical partitioning of the desired system in work packages is insufficient to cover many multi-disciplinary aspects related to the sensitivity of the instrument as a whole.

The second aspect is the System Health monitoring. The LOFAR instrument is in it’s nature similarly redundant as biological sensing systems such as the skin, the retina and the hearing sensor. However there are so many components for sensing, computing, storing and transporting that at all times a fraction of them will not work. Though the redundancy enables a graceful degradation, an understanding and possible anticipation of the degrading is essential to guarantee sufficient operation over longer periods of time as well as impact on the simultaneously running observation processes. Moreover initial heterogeneity of technologies and additional repairs and upgrades across the instrument will greatly complicate a coherent modeling of the instruments behavior from the bottom-up.

**System modeling**

There are many different views on the instrument; however the core of it’s behavior is expressed in a so-called Measurement Equation. This is a complex matrix equation describing the transformations of source signals along the entire signal path. The first sequence of transformations (i.e. integration of the source signals across the sky, the ionosphere) is approximately inverted by the instrument transformations, so as to reveal the original signal sources from the integrated signals received by the instruments. Essential is the self-calibration using known sources and the sky to calibrate the instrument: the before-mentioned inversion requires solving/calibrating the equations modeling the signal transformations, i.e. the ionosphere model and the beam-shape of the antenna and station beams with the acquired data. The behavior of the antenna, the LNAs, the clocks, the digital filters etc, is modeled from the physical principles. Simulations and measurements refine the model so that the behavior can be associated with component design parameters, and allow for an optimization of the design parameters against the requirements. However the relationships between the astronomical requirements and the specifications of the individual components are hard to find. A few years of preliminary design study and discussions at system level reveal that a common model to integrate all the effects. To study the impact of local deviations such a model would have been ideal since the interactions of the different functional steps is complicated to understand. The models and measures used in the various disciplines often seem to be incommensurable and construction of an overall consistent and coherent model has not been feasible within a limited amount of time. It may turn out that emergent behavior will result truly after full construction.
A subsystem solution approach: the LOFAR health management

The LOFAR System Health System is described in [Cabot, 2005]. The LSHM maximizes the system uptime and depends on the principle of graceful degradation per failing component. The LSHM depends on data that is automatically generated by the LOFAR instrument. The focus is on deviations from normative behavior (symptoms). The function of the LSHM is primarily diagnostic, it enlarges diagnostic accuracy. The LSHM is aiming not to be a rule-based system, i.e. symptoms and causes are not to be explicitly manually defined. Rather a model-based design is pursued, i.e. the deviations from the normative system behavior are monitored. The basis for the structure of the model of normative behavior is to come from the physical component structure, i.e. it is a modular composite model of the first-principles models of the components to be analyzed. A flexible approach is pursued such that the progressive design of subsystems can be supported as well as future repairs and upgrades. The models of components and subsystems are an essential part, since the overall system behavioral model is synthesized from these models; this model is calibrated by actual monitoring data.

A non-intrusive so-called fault detection is done using the deviations. In a so-called isolation step, reconfiguration and local self-tests can provide additional diagnostic information. The state of affairs for the prototypes is the diagnostic system for the prototype stations. This is based on subband monitoring data and snapshots antenna-correlation matrices. One distinguishes different pre-decided classes (zeros, abnormal, no signal, low signal, medium signal, high signal) based on the power in the bands and in the correlations. From expert knowledge alone some relations are made to circumstances and conditions of components to provide at least limited guidance to the fault detection process. Health variables are parameters of behavioral relationships between observables. The fault detection is the non-intrusive part of the diagnosis, and it is to be the computationally easy phase, whereas in the fault isolation one considers Lydia-based diagnostic finite-domain solvers [Pietersma, 2005], i.e. a constraint satisfaction procedure. This aims to find the health vectors consistent with the measurements. These health vectors have to relate them to the symptoms causing the issues. This relationship is crucial, but not a priori apparent. The behavioral relationships are a priori synthesized from the subsystem models. A combination of a finite set of diagnosis and a priori stated relationships cannot provide early abnormality detection.. The chosen Constraint Satisfaction Programming approach (CSP) approach leaves room for model-driven improvement as it relates to models of the processing architecture rather than to the semantics of the signal paths from the global results of that processing. The initial detection modeling, based on subsystem models may in the future evolve to include the Measurement Equation, quality of the sky image and blind projection methods already applied for mitigation of unwanted signals. The resulting global models can be more effective for monitoring the coherency of the signal processing.

5.2.4 A refinement of the problem domain

System modeling

A model is derived from first principles corresponding to the desired behavior. The global application requirements are addressed assigning requirements per sub-system, but some global requirements can only be verified by considering the application and it’s implementation in the system as a whole. Even if there is a model of the underlying design the system behavior cannot be related one-to-one with this simplified model, because unknown influences affect the system. A complete a priori model of the system behavior cannot yet be construed.
Summary of the phenomena

Time-related disturbances are a common problem. Quality and performance requirements increase the need to optimize below a certain "nois-floor". Consequently a new type of system and model generated disturbances demands attention; these are time-related disturbances caused by higher-order dependencies in the system that could be ignored before.

Temporal variations are found in the environment as well as in the system itself. In the hot-strip mill the environment includes temperature and humidity while different types and qualities of steel are being processed under different circumstances. Moreover the system itself is subject to wear and tear. In service networks, temporal variations are also found within the system, because the number of nodes, their connectivity and their software installation is constantly updated. Moreover the nodes themselves are changing over time: wear and hard disks getting full. The environment of service providers has become increasingly harsh. In a decade the number of users has decreased exponentially, while hackers and viruses make for uncontrolled effects. In sensory networks variations come from the environment, such as the weather and an increasingly hostile radio frequency band; another issue results from operating modes and schedules of the instrument as multiple observations can take place in parallel. In the three cases temporal variations occur due to varying conditions and operational modes.

Global disturbance actually appears in these real-world cases. There are disturbances that cannot be locally mitigated effectively. Hence, conventional detection strategies leave room for improvement. Fault and disturbance propagation is very complex, only a complete model of the underlying principles allows for a diagnosis. Attempts to construct system behavioral models from subsystem and component models failt, disturbances cannot be explained by the first-principles model, and cannot be modeled as a surplus of a nominal model.

Differences in setting of FDI between conventional systems and LADS

Detection of structural changes poses little a challenge, when (a) the physical principles of an information source are fully known, observable and understood, (b) disturbance test statistics arise from algebraic manipulation and analysis, and (c) model parameters relate one-to-one with assumed process coefficients. There are apparently some aspects of networked applications and the supporting distributed systems suggesting that the applicability of conventional approaches needs to be reconsidered. An overview of the differences between direct controlled systems in FDI and LADS indicates the possible causes of the limitations of conventional methods for LADS. These differences are shown in Table 5.2.

<table>
<thead>
<tr>
<th></th>
<th>Conventional</th>
<th>Local Autonomous Distributed Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>known faults/disturbances and causal relations</td>
<td>unknown disturbance propagation</td>
<td></td>
</tr>
<tr>
<td>a single function or a function per module</td>
<td>global unseparable function and qualities</td>
<td></td>
</tr>
<tr>
<td>single environment and control stimuli</td>
<td>common as well as local conditions/control</td>
<td></td>
</tr>
<tr>
<td>central controller, central direct control</td>
<td>locally autonomous, hierarchical control</td>
<td></td>
</tr>
<tr>
<td>requirements/conditions frozen design phase</td>
<td>evolution of requirements and conditions</td>
<td></td>
</tr>
<tr>
<td>closed system</td>
<td>open/evolving systems</td>
<td></td>
</tr>
</tbody>
</table>
5.3 Analysis of possible causes

In this section we analyze the origin of the disturbances, and we analyze the causes of the limitations to detect these disturbances by conventional strategies.

5.3.1 Control strategies are inadequate

Modeling is essential for control as well as detection

Modeling is crucial in all three cases for achieving global objectives. We find similarities in all three cases considering the reason for modeling, the control strategy, the modeling approach, and the simplifications and assumptions being made. System modeling is required to support control and verification for the system operation in achieving global objectives, i.e. the model serves to determine appropriate control stimuli as well as to verify through detection whether the system remains within acceptable operating parameters. In case of the hot-strip mill a physical-principle model complemented with statistical models is used to estimate appropriate local-set-points per mill for the forces to achieve the desired thickness at the end of the milling street. In service networks, implicit models are used to allow for reliable communication between servers on the highest levels of the protocol stack (e.g. OSI). Moreover shared models are used to determine the location of information and remote procedures. These models are logical data and computing models by nature. Models of a physical phenomenon under observation are studied in sensor networks. The environment as well as the signal processing instruments is required to control the instrument and route the collected data; hence their models are essential to interpret the recorded data, but also serve to verify the correct operation. The control strategy in such cases is that of local autonomous control, the underlying assumption being that global objectives can be met when local requirements are fulfilled. In the hot-strip mill each mill is given an initial set-point to the local controller for decreasing an expected input thickness to a target output thickness. In a service network each node performs its operations autonomously, running its local program to fulfill its local performance requirements. In a sensor network local operation of sub-systems is centrally scheduled but, the operation is locally autonomous such that processing can continue even when control is absent.

A model is never complete

The control and monitoring approaches in distributed, locally autonomous processing relies much on strict local control from models, based on a priori domain knowledge. Control stimuli may be locally optimal, but the absence of a global behavioral system model prevents optimization of overall system control for global objectives. The detailed a priori modeling from physical or logical domain principles is apparently not adequate to practically deal with the many variations in conditions and configurations. Particularly it is very easy drowning in the details of a local disturbance while failing to capture the system dynamics, as relevant to overall system objectives. We immediately admit the great benefits of domain and expert knowledge preferably in the form of exact models of the system behavior: if present.... use it. The different scales and abstractions, heterogeneity of technologies implies that an exact measurement equation and detailed behavior of components cannot be coupled into a single coherent and consistent, analytical or simulation model. Apparently the attempt to exactly model distributed systems for the purpose of control and monitoring does not achieve to express the desirable behavior such that relevant changes in the system behavior can be detected early enough.
Consistent and coherent models of the system behavior are not easily obtained, and the propagation of disturbances is harder to prevent. This shows that there will always be gaps in an a priori model of the system as well as in cause-and-effect models.

... therefore detection is required

"Why do we need detection and accommodation at all?". As a consequence of the incompleteness of an a priori model, discrepancies occur after control actions, and are left to the detection and accommodation. The answer to the question is that not all the desired behavior of the system can be enforced from an a priori model of the system and its environment, because the nominal process model is incomplete.

Key remark 5.1: need for detection and accommodation
Detection and accommodation are needed to compensate for disturbances resulting from limitations of direct control based on the nominal process model.

When all the required control stimuli can be generated from the nominal process model, there can be no disturbances but only faults and failures that are identified within the model. However disturbances occur; hence the changes in the system and its environment cannot be all detected and accommodated in practice, while FDIA proofs the possibility to improve on the nominal process model(s).

Detection and accommodation are part of the system

Since FDIA is applied to detect and accommodate disturbances and faults, the performance of a system can no longer be derived from the performance of the nominal process model used for direct control. Hence, without further argument we state

Key remark 5.2: performance of system includes detection and accommodation
The performance of the system is performance of the system's direct control derived from a nominal process model combined with the performance of detection and accommodation correcting the systems control.

The objectives of modeling for detection and accommodation can be derived from this key remark. Despite the risk of over-emphasizing this, we identify what is to be actually detected: the limitations of the system model applied for control cause the disturbances to be detected and accommodated, the model limits understanding of causes of faults that need to be prevented. This suggest to reconsider relation between a nominal process model and the required detection model. The disturbances to be detected and anticipated require a model for detection and accommodation that is complementary to the process model applied for control. Particularly this complementary model needs to be susceptible to global disturbances that are beyond the scope of a priori first-principle local models.

5.3.2 Disturbances: global disturbances

Systems are increasingly used beyond there originally specified domain of operation

Distributed systems evolve within an environment. Their behavior is characterized by independent processes that are governed by similar dynamic principles or even the same source. The assumptions, that underlying conventional approaches, do not hold outside the stable equilibria (within acceptable range w.r.t. design objectives) and complexity hits back hard!
The critical drivers from new and increasingly tight application requirements are found in all three cases. In the hot-strip mill there is a need to increase the production rate as well as the quality of the pressed steel, i.e. ever lower thickness variance. Reliability, availability and maintainability are directly related in service as well as sensory networks to customer satisfaction and operational cost, unavailability is increasingly less acceptable. In other words, quality of service is a main driver. Another aspect to mention here is the evolutionary and dynamical aspect of applications and systems. In the operational phase of systems they are used in unforeseen ways and requirements pop-up that have not been anticipated. This is even the case for systems that have been around for decades. They are pushed, through modifications, to meet new user expectations over and over again, the Westerbork observatory is a typical example.

**Dependencies between autonomous processes w.r.t. global objectives**

Conventional detection and accommodation is based on a priori knowledge rather than on data monitored by the dynamics and dependencies among disturbances, because: 1) they are to be irrelevant to achieve local control objectives; 2) they are too complex to model in a coherent and monolithic way. The rationale is that what is not modeled in the system model appears as a disturbance. Hence the challenge is actually in the modeling for detection since dependencies and dynamics of the disturbances throughout the system are relevant to achieve performance improvements. The heart of the detection problem for distributed systems is found in the interaction of autonomous processes that are not modeled if the system model is composed of independent state-space models per autonomous node, i.e. the process state-vector interactions are ignored:

**Key remark 5.3: global disturbances**

Global system disturbances and faults are generated from dependencies between distributed processes that are locally autonomous.

We conclude that distributed processing systems are too complex to provide a global monolithic system model. The achievement of system objectives is necessarily dependent on isolating some processes in the system that set some local control objectives, while trusting on locally autonomous processes. Consider the meaning of local autonomous processing: governing itself by local law (auto=self, nomos=law), or absence of direct control from a central controller. The local controller pursues local objectives (local control targets) relying on local models, while the local control objectives themselves are derived from a global model. System disturbances will not occur, when the dissemination of global objectives through a system model to local objectives is adequate. However, the global system model is a hierarchical/modular composition of simplified local processing models, while the interaction between processes is not modeled, except for communication through input/output of the local processes. The internal state variables of local processes are considered independent. There is no global system state space model covering the interaction of local processes through whatever medium. If the local processes are indeed independent, then the hierarchical composition of a global model will be valid and local disturbances will not be relevant on a global system level. Though the local control is adequate with respect to local objectives, there can be gaps in the system model causing inability to describing all relevant dependencies within the system.

Key remark 5.3 is also based on the observation that system disturbances and faults can emerge due to a limited system model for control. However, we should consider whether such process dependencies are likely. Interaction between processes exists because of several aspects:
• local controllers respond to each other through deviations from earlier processes;
• local processes share an environment;
• processes are composed of similar components, and respond in similar ways.

Are the dependencies between the disturbances of "autonomous" processes relevant in the achievement of global system objectives? The interaction between the "autonomous" processes may be dormant. However, the dynamics of these interactions are the only possible indicators for failures of the system that cannot be detected locally. A 1st-order system model is composed from simple models of locally controlled processes. Such a model does not take the interactions between the processes into account. The global disturbances appear as a result of local disturbances, assuming that the system behavior is indeed the composition of the behavior of local processes. Local disturbances can only result from limitations of the local control, say of the local model. If the causes of local disturbances are all local (i.e., no interaction between processes), there will be no room for improvement in the detection of global disturbances and failures from a global model. If the local control can keep the local processing in equilibrium by fulfilling the local control objectives, and if this also implies the fulfillment of global control objectives, then the local errors will not likely accumulate to system failures. However, we see in practice that they do. Variations in the hot-strip mill cause undesirable thickness variations while an improvement using cross-mill variables was possible. The acceptable use and performance of service networks cannot be adequately enforced using local direct monitoring and control only.

### 5.3.3 The complexity of modeling

**The system behavior is not as predictive and invariant as designed**

"Expressing system requirements involves specifying against unwanted behavior in responses to unforeseen events while many applications are now targeting environments that cannot be considered as closed and for which knowledge representations will necessarily be incomplete", [Lisboa, 2001]. A system can behave unpredictable because of its evolution as well as due to the partially unpredictable environment it is part of.

We have encountered both dynamics in requirements and system conditions, as well as changing and diverse system use for applications with increasing demands on performance and quality. Conventionally the system is developed with a single application in mind relying on known physical and logical principles for a one-time design phase; however usage and conditions may change over time. The systems in conventional FDI are monolithic machines with tightly coupled actuators and sensors relying on direct feedback control or a priori determined schedules for batch processing. The complexity of networks is found in the geometrical spread of processing, the collaboration of locally autonomous processing, the diversity within the system, the scale and size of the systems, and the automated batch-oriented processing i.e. having to deal with operating modes and conditions. Disturbances in distributed systems are complex due to dependencies across processes, affecting the global system performance.

Conventional modeling is a 2-tier approach, as the assumed physical or logical principles fixate the design of the detection model. The system behavior, encountered in an automated industrial plant, may be thought of as a designed system of coupled processes, but usually the system behavior is much more complex than the "blue print". Controlling such a system is only
possible within a small manifold of the state-space, since outside the manifold the assumptions of independence and linearity are false.

**The systems and sources of disturbances are not independent**

The systems behavior and the disturbances is more complex in LADS than in the closed and monolithic control and communication systems. Firstly, in the processing of the reference cases complexity is found in: 1) the large number of independent processes, 2) the autonomy of the processes; 3) the distributed collaborative aspect. In the hot-strip mill a cascade of mills is locally controlled to achieve the global plate thickness. In service networks, servers are locally independent yet they can only provide information services together.

The scale contributes to the complexity of a system in terms of the number of sensors and actuators, the number of i/o's, the number of samples taken per time interval, and the distributedness of the system due to the physical size. With a large scale the need for distributed sensing, acting and processing together with a supportive infrastructure arises. In a hot-strip mill there are 1 to 20 mills distributed over 10-100s meters, while service networks consists of 10s to 1000s of participating servers, that can equally easy be located in one room or on one planet. In sensory networks there is a huge variety. LOFAR consists of 10000s of sensors and processors distributed over 100s kilometers. The complexity of the disturbances in the distributed systems is higher than in monolithic systems, particularly considering the nature, dependency and impact of dynamic variations. Local dynamic variations cannot be considered in isolation, because they arise from interacting processes that are physically or logically separated. The disturbances are not independent as the processes share common factors.

In the hot-strip mill the processes communicate implicitly through the global (control) objective, the plates that are exchanged, the environment and the similarities in the process of wear and tear per mill. In the service networks, the logical processes and processing are mapped onto a shared infrastructure called network, the processing nodes in the network, the power supplies etc.. Moreover the software configurations may also be shared. Consequently attacks, bugs and faults are not limited to local nodes. In LOFAR the environment is locally similar, the physical phenomena studies are globally similar, and the local distributed processing systems are also similar. Coherency in different signals paths enables a correlator to amplify sensitivity for celestial objects, but similarities in the disturbances cause degradation of the astronomical observation!

**Simplifications of the modeling as a common strategy**

In the cases we see several similarities in the modeling approach considering the desire to apply homogeneous conventional and monolithic models that are well-understood, starting from known physical or logical principles. Tight quality and performance raises the need to reach deeper into the “noise”. Looking at a finer grain while pushing performance the assumptions that allowed for simplifications start to break apart. We need to consider complex dynamics and dependencies which seemed irrelevant before. In the cases, the state-of-practice is the preference of homogeneous conventional monolithic models that are well-understood using physical or logical principles. In the hot-strip mill the starting point is the physical force prediction model per mill, disregarding most complexity that belongs to the system. The physics are understood; hence it is considered reliable though not adequate. In the service networks each protocol layer and communication model considers particular aspects to provide an abstraction to the higher layers. These models are used locally on the nodes to handle incoming
requests; global models for the behavior of service networks are not used for local execution. In the sensor network various architectural views are used in the design and operational phase. They provide different models of the instrument that are only coarsely coupled. Even for the instrument the particular subsystems and domain-experts rely on highly abstract processing models, so that each focuses on the signal processing quality aspects in relation to a detailed mono-disciplinary behavioural model. There is only a weak coupling of the models for RF signal processing, digital processing, imaging and astrophysical system requirements. If the models need to be monolithic, well understood and derived from physical or logical principles etc, the models come as small building blocks. The modeling of a system then faces two challenges: (i) how can the validity of the building blocks be guaranteed; and ii) can global requirements be fulfilled. The modeling approaches rely on simplifications to deal with the complexity of the system: separation into local concerns, hierarchical and modular construction and linearization. These strategies rely on some assumptions to simplify the problem of system modeling:

- global requirements can be met if local requirements are met, i.e. the control strategy enables a particular modeling approach. We can refine this into assumptions 1) the composition of local processes requirements can be derived given global requirements and 2) local control can keep local process states in an equilibrium meeting local objectives.

- process interactions beyond the input-output relations can be ignored. If a local process fulfils its requirements, then marginal variations will not affect other processes significantly. Specifically the processes can be required to be BIBO-stable or even LTI to meet this assumption. Local control and scheduled maintenance are designed to ensure a local equilibrium which implies time-invariance and linearity.

- only a limited and known part of the state-space has to be considered. With this assumption the various conditions and operating modes are covered, given the conditions and mode. A particular local model is chosen to be valid within a certain regime; within this regime behavior is assumed to be quasi-stationary.

The validity of these assumptions is at least questionable. The modeling of the dynamics and their dependencies shows in the case of the hot-strip mill and of the service network the existence of a structure that can improve the performance of the systems. In either case, performance and RAMS requirements have triggered the need to consider some higher-order dynamics and dependencies within the system. Similarly the requirement of sensitivity in astronomical instrumentation, while man-made radio interference increases, implies that the instrumental effect must be modeled and calibrated for system noise to decrease.

The common control strategy is to consider the distributed processes to be locally autonomous; the global objectives are delegated by deriving local processing objectives. This is not the case in conventional systems, where the control is central. Further we see attempts to compose a system model for control and detection from known local behavior. This is an attempt to apply conventional modeling techniques. There are assumptions and simplifications necessary to allow for hierarchical and modular composition of a system model from linearized models of the processes within the system. Consider in particular the enormous state-space complexity for distributed systems if all interdependencies and non-linearities are modeled. The assumptions typically reduce the complexity to a single state-space vector and one state-machine or transition matrix per autonomous node. These assumptions are not necessary in case the sys-
tem is monolithic. In order to achieve desired quality it appears to be necessary to consider the higher-order dependencies, beyond the simplifications in the nominal process modeling.

The limitations are caused by the conventional objective to minimize model complexity. These limitations cause a loss of information on the structure in residuals when they are projected to the model parameter space, as the model can become invalid when the modeled process changes. Since the behavior of complex systems is hardly explained by the "blue print", the acceptable functional behavior should be distinguished from the observed behavior. Conventional approaches fail to distinguish acceptable uncertainty from suspicious patterns/trends. Even when physically plausible models can be obtained, there is no guarantee that these models are suitable for detection of the patterns that indicate a potential undesirable trend. The structure of the system is not time-invariant, which makes the model topology unsuitable for parameter-based signature computation.

Complexity management relies on a divide-and-conquer approach (monolithic modeling fails). Modular modeling and control requires the assumption of locally independent processes, stationary behavior and linearization. Yet global dynamic models offer room for improvement to validate correct global behavior of the systems.

**The patchwork strategy**

In any approach to detection, some kind of model extension is used in the signature computation and decision making to catch the disturbances. These extensions are superfluous for the nominal process model (Figure 5.1). However, redundancy is crucial to achieve sensitivity for detection. We have seen in section 4.4 that the redundancy can be analytical, expressed in additional functional relation between observed variables or by statistical metrics on the parameters or left as residual of a model. The design of this redundancy depends on the type of problem and application. In conventional applications, the nominal process model itself is preferably exact or at least has a statistically optimal model order to prevent inaccuracy in the model parameters and modeling artifacts. Then redundancy is designed by explicitly expressing abnormal system behavior in fault models of physical processes by coefficients or residuals of the nominal process model.

This approach is suitable for separating variations under normal conditions from variations due to disturbances. It can be used for the prediction and prevention of faults, unless the nominal system model parameters or it's residuals are incapable of reflecting changes in the system. The overall validity to describe the inherent structure of the system is critical. A complexity reduction of the system model is required to deal with the "curse of dimensionality" and the limited human capacity to interpret models. However this yields a bottleneck when emerging faults are to be detected as a result of changes. Since the compositional model has significant gaps in the state space representation, such a nominal system model is inadequate as a starting point of detection of global disturbances.
Figure 5.4: In conventional methods nominal process models are exact, redundancy is external.

Simplifications to the modeling of instances and abnormalities

Firstly only a finite number of conditions and operating modes are considered in the cases; all other behavior is not considered. In the hot-strip mill the conditions are predefined classes of steel composition, thickness target etc. The logical models for controlling and monitoring the service networks is based on finite-state models, covering only known states of correct behavior and some failure modes related to network resource faults. There is nothing in between. Secondly just a limited number of disturbance types are taken into account. The distribution of the residuals of process-coefficients associated with correct and faulty operations of the system is assumed to be known in advance. In the hot-strip mill faults that cause the system to fail require immediate repair, while other disturbances may be polished away locally. The dynamics of temporal variations are not considered. Similarly the behavior of a sensor system is characterized by a large and hierarchical finite-state machine. Considering the closed model of the normal system behavior and the closed model of the fault states, we conclude that a finite-state model describes the disturbances and faults where no new states are added during use nor the transient behavior between states is considered. The detection mechanism essentially becomes a classifier, where the disturbance distributions are assumed to be known a priori.

Ideally optimal control leaves a stationary residual, an error without structure. However, the system and the system model are no longer monolithic causing a delayed effect of control actions, and acceptable variations prevent stationary residuals even for a global optimized solution. Moreover if control is locally optimal it is not necessarily globally optimal due to a lack of complete system control. The dissemination of global to local control objectives as well as the composition of a system model from simplifications often assumes linearity and process independence. The composition of such a system model ignores the dependencies and common factors underlying multiple processes within the system that are likely to cause system disturbances and dynamics, resulting in faults. These disturbances and dynamics are not isolated in any parameter of the system model or system residual, since the composite system model lacks parameters for interdependencies. Thus the system model by hierarchical and
modular composition is less suitable as a reference for signature computation. Recall the modeling artifacts we describe in section 2.3.6 as a cause of disturbances.

Signature computation is conventionally based on parameters of a process model when the process is assumed to be known by design; otherwise signature computation is based on the residual. Acceptable boundaries on physical properties, derived from model parameters, can be determined in case one assumes a controlled process, whereby the control of the process aims to keep the process within an optimal equilibrium. If control guarantees local and global stationarity, then normality is equivalent to a stationary residual, and any local non-stationarity implies abnormality. The common case is to isolate specifically known features from known fault signatures or physical limitations by matching filters or extraction of physical properties; otherwise the general approach is to compute signatures that are universal but robust to noise. The Eigen structure of the residuals allows distinguishing the null-space from the signal space. Similar projections can be achieved by analysis filters that ignore particular components in the residual. A boundary test detects a relevant signal component when it emerges after projection to the modeled process. The signature projection to isolate disturbances is feasible if the process equilibrium is preserved guaranteeing stationary independent errors through time.

**Global statistical methods revealed patterns allowing for improvements**

Literature and experiments reveal that the dynamics and dependencies of disturbances allow for data-driven behavioral modeling that can help to anticipate and prevent serious performance degradation. This indicates the potential for a more holistic approach to modeling. In the case of the hot-strip mill we have shown an improvement of the role-force prediction through the use of a neural network, where the neural model is initially derived from data and continues to adapts showing sufficiently stable learning. Global analysis of the dynamics in the service network reveals patterns in the session traces; the statistical analysis of these patterns is possible with a data model of the network traffic. Surprisingly, the model has not been inspired from logical principles but the classes can be retro-associated to services usage or misuse.

**5.3.4 Pitfalls of conventional approaches**

In this section we show that the conditions for conventional methods lead to limitations. It is no longer clear how individual processes contribute to the quality of the total application. The cases illustrate that a first-principles model no longer provides the required understanding. Time-related dependencies are relevant to achieve the desired quality and global requirements need consideration of the application as a whole in the system modeling for the purpose of detection.

**Compositionality**

Simplifications allow for modeling a complex reality, but do not provide a robust and reliable method for monitoring the complex distributed sensing, processing and governing networks that dominate our automated industry and networked society. Only a specific fault model allows a specific detection of and accommodation to that fault. There are many different types of disturbances and faults but they are all rare. In cases where the physical plausible model is absent, one has to rely on highly abstract, global models derived by statistical analysis. Such models do not allow for a quick response.
Key remark 5.4: Control-driven system models fail to explain disturbances

The true complexity emerges as disturbances that cannot be explained from a composite model. Disturbances appear as a result of decreasing performance of the overall system control, which in turn is the result of incompleteness of the system nominal process model, i.e. unforeseen system behavior.

The simplifications required to compose a system model yield structural deviations both in the local as well as in the interoperating processes. These accumulate into intertwined modeling artifacts that cannot be observed in isolation; model validity can only be ensured in a very limited part of the state space. Overall we conclude that hierarchically and modular system models composed through simplifications only describe the tip of the iceberg concerning the dynamics of the system as a whole. The appearance of time-related disturbances as a result of dependencies between processes indicates that such a model cannot be used to explain or detect the structure in such disturbances. Control-oriented modeling of distributed locally autonomous systems constrains the overall system state transition within regimes where independence of processes and disturbances is approximately achieved. This allows for a composition of a system model based on models of local processing behavior.

Due to the complexity of the system and its disturbances a model of high granularity cannot be found from physical or logical principles to explain the interactions between the independent processes. The accuracy of the models is not improved using mathematical analysis, but data-driven models are immediately successful in capturing the dynamics. Exact models are usually valid under limited conditions. The many variations in configuration and conditions actually raise the need for taking many exceptions into account. In case of distributed locally autonomous systems, mathematical exactness of the model is hard to achieve. Artifacts resulting from control-stimuli that are structurally off for a globally optimal set-points are not computed.

Superposition

The patchwork strategy starts from exact or “statistically optimal” nominal models and the disturbances and abnormalities are considered a surplus on top of the behavior explained by such a nominal model. The abnormalities are viewed upon as being superposed on the normal system; consequently the model for detection is the nominal model with some patches to describe the abnormalities on top of it from this conventional point of view. This view is actually too limited, since:

- artifacts arise that cannot be explained in this fashion
- disturbances cannot be traced to variations in the parameters of the model
- architecture of the whole model is not correct; hence the parameters do not explain changes in system behavior.

Determinism

Only a finite set of disturbances and faults are considered by conventional detection approaches. However in the real-world cases we see:

- there are too many different types of disturbances and faults to be described a priori;
- future conditions and configurations cannot be anticipated;
- transient behavior is ignored, but the system changes intrinsically, i.e. system behavior is not modeled by the nominal process model superimposed with a model of disturbances.
The distribution of faults and disturbances cannot be assumed in advance, because there are so many exceptions that it is not feasible to account for all of them. Moreover future conditions and configurations of the system cannot be anticipated, even less so for the future behavior of the system as it evolves in an environment that changes. The impact of the absence of a priori knowledge on the application of conventional FDI approaches is critical.

**A priori optimization is impossible in real-world cases**

Some fundamental questions are raised when we are accepting that unknown states exist and state transitions occur, anticipating a system behavior that is more complex than what is a priori captured through a hierarchical system model. When the classes of normal and abnormal behavior are no longer closed sets, there is no way to determine reliability, sensitivity and promptness of the detection and accommodation. In fact we can no longer consider the detector to be a classifier. The challenge to be addressed in chapter 6 is deriving a consistent set of requirements on modeling and signature computation for these conditions.

### 5.4 Problem statement

The dynamics of local disturbances need to be analyzed in the context of both local and global system performance to anticipate failure. This requires the analysis of the dynamic interactions between the disturbances of "autonomous" processes. Disturbances and process dependencies need to be analyzed on a system scale; failure to achieve objectives can only be anticipated if the system is considered as a whole; interactions may be dormant when local control achieves local objectives but otherwise accumulate into system disturbances.

**Challenging characteristics of LADS**

A major challenge in the detection for LADS is the absence of a monolithic "exact" model for the system behavior. This results from the absence of known underlying principles that govern the interactions in the system and with the environment. The systems are composed of modules that are understood but the global behavior has not been modeled to an adequate level of detail. There are huge amounts of data available but there is no adequate overall model. Common hidden features (i.e. unknown state-variables) between different processes cause time-related effects and interaction between processes that should behave independently. Hierarchical and modular composition of systems and of system models from modules that are well understood does not yield an overall system or system model that explains the behavior adequately. System theoretical models and probabilistic models are too rigid or generic resp. and fail to capture the globally coherent dynamics associated with the desirable operation of the system as a whole. The system changes are likely to invalidate a nominal process model for detection.

**Problem statement**

We conclude that the intersection of desirable behavior of networked applications and the actual behavior of distributed locally autonomous processing systems that interact to provide global performance is not easily modeled from logical or physical principles. Yet this intersection is essential to derive control stimuli as well as a reference model for early detection of unforeseen changes.

**Detection Problem in LADS:** The prevention of harmful failures in LADS depends on the development of dynamic models for global system behavior that
allow to separate acceptable from potentially harmful dynamics assuming inter-
dependence between distributed processes.

This is a problem due to the complexity of distributed systems and the consequential invisibility of disturbances. The dormant disturbances accumulating into faults and failures will go undetected by a conventional FDI approach. They will become observable only as they fail. In any case where such emerging system changes are present, they need to be isolated from artifacts that arise from limitations of a nominal process model. To prevent the accumulation of disturbances into inseparable faults, the interdependence needs to be observable in a detection model. The modeling requirements are different for the purposes of a) expressing the desired and intended behavior; b) control under normal operating conditions, and c) detection of unforeseen disturbances that may lead to faults. The combined requirements overconstrain a single model, if it is to be physically plausible or statistically optimal.

We come back to the questions we have started with at the beginning of this chapter. What is the challenge? The challenge is modeling for detection despite the complexity of system behavior and abnormalities. What are the challenging disturbances? Global disturbances resulting from abnormalities that are intertwined with the system! These cause the system to deviate rather than to superpose effects. What is the origin of these disturbances? They result from the complex interaction between processes inside the system that are considered independent, and from the interaction of the system and it’s environment. Why are conventional methods insufficient? Conventional methods have modeling limitations. Particularly modeling simplifications are necessary to deal with the complexity explosion yield a model inadequate for detection of global disturbances.

5.5 Conclusions

In the real-world cases presented in section 5.2, there are global disturbances resulting from abnormalities that are intertwined with the system. The origin of these disturbances is the complex interaction between processes inside the system that are considered independent (key remark 5.2) but are influenced and linked by external phenomena and shared resources,. Conventional methods have modeling limitations. Particularly modeling simplifications are necessary to deal with the complexity explosion, since classical system differ fundamentally from locally autonomous distributed systems, table 5.2. We introduce the issue of complexity vs. modeling capability in chapter 4, classifying existing detection methods by the complexity of systems and abnormalities. Now this issue is much clearer in the domain of locally autonomous distributed systems with global functions: the global disturbances are due to abnormalities and the system being intertwined (key remark 5.3). The nominal process model fundamentally limits observability for the system-abnormality interaction, which is conventionally a reductionistically simplified system model from the first principles or a statistically optimal probabilistic model, key remark 5.4. Observability for global disturbances is not achieved by patching fault sensitive models to such nominal models, as shown in figure 5.4. The challenge is in modeling for early detection despite the complexity of system behavior and abnormalities. The new understanding of the problems provide a good starting point to pinpoint at what modeling capabilities are needed for early detection in the next chapter.