Datenverarbeitung in der Industrie 4.0

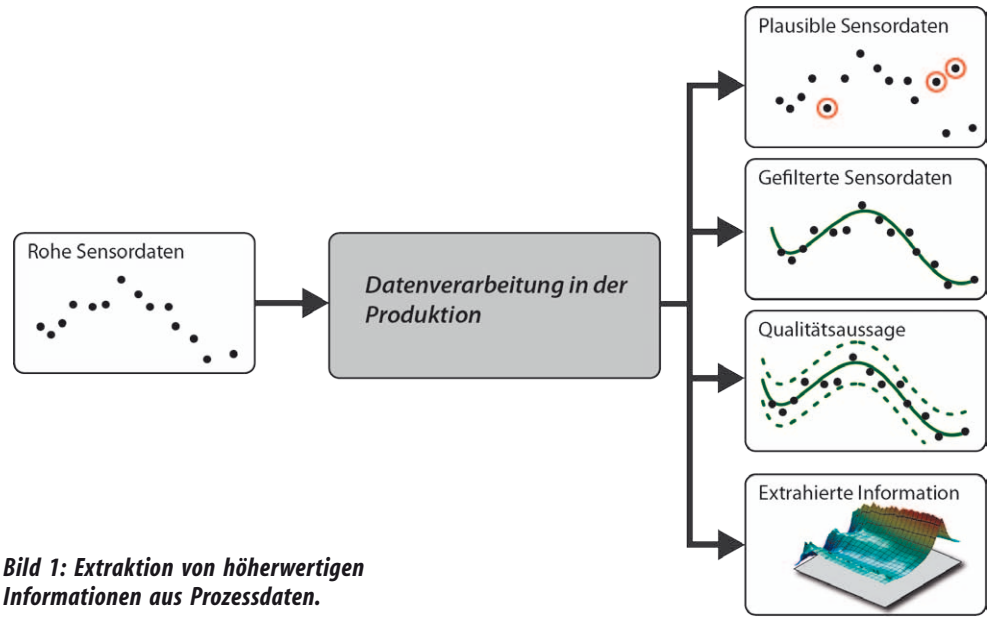

Bild 1: Extraktion von höherwertigen Informationen aus Prozessdaten.

In den letzten Jahren wird zunehmend der Begriff Industrie 4.0 verwendet, welcher die vierte industrielle Revolution beschreibt. Durch die Miniaturisierung von Sensoren und die Entwicklung neuer Übertragungstechniken wird die moderne Informations- und Kommunikationstechnologie immer mehr mit der Produktionstechnik verschmelzen. Die Basis dabei ist, dass alle Informationen in Echtzeit für den Produktionsprozess verfügbar sind, und von den einzelnen Systemen verwendet werden. Diese Systeme werden Cyber-Physical-Systems genannt, welche die digitale Welt mit der physischen Welt vernetzen. Dies führt dazu, dass alle Beteiligten im gesamten Wertschöpfungsprozess, d.h. Menschen, Objekte und Systeme, miteinander vernetzt sind. Ein wichtiger Bestandteil sind dabei intelligente Sensor-Aktor-Netze, die zum einen mittels räumlich verteilten Sensoren die Umgebung erfassen, und zum anderen mit ihren Aktoren auf die Vorgänge einwirken können. Durch diesen Zusammenschluss entstehen dynamische, echtzeitoptimierende, und selbst organisierende Wertschöpfungsnetze.

Der vom VDMA heraus gebrachte Leitfaden zur Industrie 4.0 soll dazu dienen in einem systematischen Prozess Potentiale zu identifizieren und daraus geeignete Ideen zu entwickeln. Dabei stehen die zwei Werkzeugkästen für Produkte und Produktion im Mittelpunkt. In diesem Artikel wird auf die beiden Schwerpunkte „Integration von Sensoren/Aktoren“ und „Datenverarbeitung in der Produktion“ eingegangen.

Herausforderungen

Die voranschreitende Vernetzung aller Komponenten in der Wertschöpfungskette lässt neue Herausforderungen für die Datenverarbeitung entstehen. Die klassischen, einfachen Verfahren und Algorithmen sind nicht in der Lage die neuen, komplexen Fragestellungen, welche sich aus der Analyse von Prozessdaten ergeben, zu beantworten. Moderne, automatische Algorithmen können zwar Prozessdaten weitestgehend automatisch analysieren und sehr komplexe Zusammenhänge erkennen, allerdings sind diese nicht in der Lage mit der stetig steigenden Datenmenge in angemessener Zeit umzugehen. Hier ist noch viel Forschungs- und Entwicklungsbedarf notwendig, um leistungsstarke Algorithmen zu entwickeln, die Prozessdaten über die gesamte Wertschöpfungskette hinweg analysieren, und automatisch versteckte Zusammenhänge erkennen können.

In der gesamten Wertschöpfungskette kommen immer mehr Sensoren und Messeinrichtungen zum Einsatz, deren Daten bei der Auswertung zwar ein enormes Potential bieten, allerdings dessen heterogener Charakter auch eine besondere Herausforderung für die Datenverarbeitung darstellt. Das Verhalten und die Charakteristik der Prozessdaten unterscheidet sich sehr stark je nach physikalischer Herkunft der Messgröße, beispielsweise Druck, Temperatur, Luftfeuchte oder Farben. Weiterhin können Prozessdaten in der Anzahl der Dimension sehr stark variieren, beispielsweise müssen die Daten von Beschleunigungssensoren, Kamerabilder und 3D-Punktwolken unterschiedlich verarbeitet und miteinander verknüpft

Chancen und Nutzen der Datenverarbeitung

Für verschiedenste Anwendungen ist das Kernthema die Verarbeitung von Daten, da damit eine kontinuierliche Erfassung von Informationen über einen (technischen) Prozess möglich ist. Zudem können durch die Verknüpfung verschiedener Informationen neuartige Erkenntnisse gewonnen werden. Die Art und Weise, wie die Verarbeitung der Daten durchgeführt wird, reicht von einfacher Dokumentation, über Prozessüberwachung, bis hin zur autonomen Prozessplanung und -steuerung.

Die daraus gewonnenen Informationen dienen der Entscheidungsunterstützung sowie der autonomen Entscheidungsfindung im Besonderen für Steuerung und Regelung, sowie Diagnose und Alarmierung. Dadurch werden die Qualität der Produkte und die Effizienz der Produktion erhöht, als auch mögliche Fehlerentwicklungen frühzeitig erkannt. Dabei kommen eine Vielzahl grundlegender Techniken aus der Statistik, dem maschinellen Lernen und Data Mining zur Anwendung. Aus der Analyse historischer Messwerte werden unbekannte Zusammenhänge offengelegt, oder Prognosen über das zukünftige Systemverhalten bestimmt. Diese gewonnen Erkenntnisse ermöglichen eine kontinuierliche Optimierung von Systemen, Prozessen, und Strategien.

Zudem können Kosten eingespart werden, da an die Maschine angepasste Wartungsintervalle bestimmt werden können. So werden nur die notwendigen Wartung durchgeführt. Darüber hinaus können schleichende Veränderung erkannt werden, auf die reagiert werden kann. Die Wahrscheinlichkeit von einem Stillstand durch plötzlichen Ausfall wird somit reduziert.

Autoren:

Felix Sawo,
Frederik Beutler
Geschäftsführer bei Knowtion

werden. Der Einsatz von RFIDs, Barcodes, und die diskreten Prozessschritte während der Wertschöpfung machen weiterhin die Unterscheidung in kontinuierliche und zeit-, wertdiskrete Prozessdaten notwendig. Hinzu kommen qualitative Unterschiede bei den eingesetzten Sensoren bzw. den Prozessgrößen: Beispielsweise müssen Größen aus einer Sichtprüfung anders bewertet werden als von einem physikalischen Sensor; auch bei den Sensoren untereinander gibt es Unterschiede in Genauigkeit und Güte, die bei der Verarbeitung und Verknüpfung berücksichtig werden müssen. Diese Heterogenität stellt eine große Herausforderung dar, mit denen zukünftige Algorithmen umgehen müssen.

Ein wichtiger Baustein zukünftiger Industrie-4.0-Anwendungen stellte die Trennung von Informationen und Standort durch den autonomen Informationsaustausch (z. B. durch Webdienste über das Internet) bei der Maschine-zu-Maschinen-Kommunikation dar. Durch diese Dezentralisierung werden die Prozessdaten nicht mehr zentral von einem Server gespeichert und verarbeitet, sondern die Datenverarbeitung findet lokal vor Ort auf den einzelnen Komponenten der Wertschöpfungskette statt. Die Verknüpfung der Prozessdaten mit den Produkten beispielsweise führt dazu dass die Produkte ihre eigene Herstellungshistorie kennen und durch entsprechende Algorithmen und Verfahren die weitere Prozessverarbeitung bestimmen können. Dies führt zu neuartigen Herausforderungen der sich die IKT-Infrastruktur und die Datenverarbeitung stellen müssen.

Eine umfassende, echtzeitbasierte, und vor allem robuste Datenverarbeitung in der Produktion erfordert neben der Erhebung verschiedener Erfassungsdaten die Plausibilisierung und Bereinigung von Störungseffekten. Im Vergleich zu anderen Messeinrichtungen in kontrollierter Umgebung unterliegen die mit Sensoren erfassten Messwerte an Produktionsanlagen starken mechanischen und thermischen Beanspruchungen. Dies führt dazu, dass die relevante Information in den Messsignalen durch Störsignale verschiedenster Art überlagert wird. Störungen wie beispielsweise Brummen, Rauschen, Messaussetzer, Sensordefekte, oder mangelnde EMV schränken die Funktionsweise und Lebensdauer der Sensoren erheblich ein. Für die automatische Prozessplanung bzw. -steuerung ist es somit zwingend erforderlich, diese auf Plausibilität zu beurteilen, und somit ein Mindestmaß an Vertrauen in die Qualität der Prozessdaten zu haben. Nur durch die frühzeitige Plausibilisierung von Sensordaten kann verhindert werden, dass offensichtlich fehlerhafte Messwerte für die Prozessplanung/-steuerung herangezogen werden. Hierfür sind neuartige, leistungsstarke Algorithmen und Verfahren notwendig.

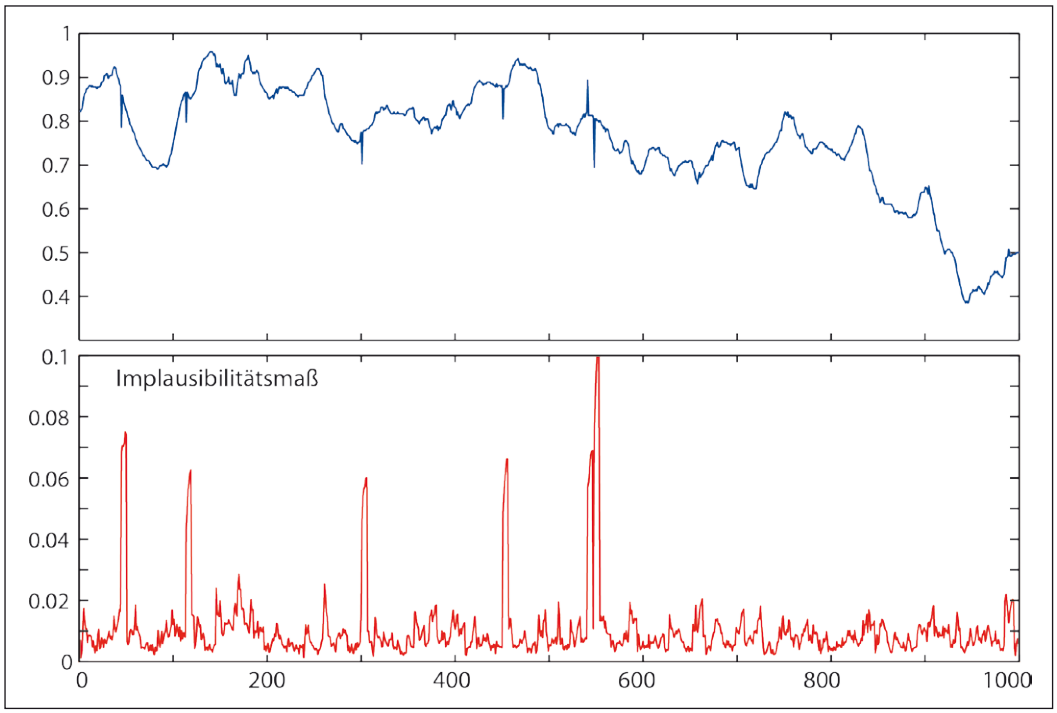

Bild 2: Darstellung der Prozessdaten (obere Kurve) und das bestimmte Implausibilitätsmaß, welches ein Maß für die Produktqualität darstellt.

Des Weiteren müssen Latenzen und Datenausfall - welche unvermeidbar bei einer drahtlosen Kommunikation im industriellen Umfeld entstehen - bei der Datenverarbeitung berücksichtigt werden. Hier bieten Algorithmen zur Interpolation und Prädiktion von Sensordaten einen interessanten Ansatzpunkt.

Verfahren und Algorithmen

Bayes'sche Schätzverfahren können zielführend zur Darstellung von Sensorinformationen bzw. Zustandsaussagen (u. a. durch Datenfusion), zur Plausibilitätssteigerung, Datenreduktion und Extraktion höherwertiger Information eingesetzt werden. Im Allgemeinen wird ein explizites mathematisches Modell angenommen, welches die physikalischen (bzw. chemischen) Eigenschaften und Zusammenhänge des zu verfolgenden Prozesses beschreibt: Beispielsweise zeitunabhängige Stoffgesetze oder analytische Lebensdauermodelle. Je genauer die Modellierung des abzubildenden physikalischen Prozesses erfolgt, desto genauer ist in Folge das Ergebnis der Zustandsdaten. Hierzu ist eine Vielzahl an zunächst unbekannten bzw. noch zu bestimmenden physikalischen und chemischen Parametern zu implementieren.

Algorithmen aus dem Bereich der Sensorfusion bieten Möglichkeiten ein Rauschen durch optimale Nutzung vorangegangener Informationen zu minimieren. Dabei werden verschiedenen Messkenngrößen, wie Abbildungsfunktionen und Rauschcharakteristika, berücksichtigt. Zudem wird bei diesem Verfahren implizit die Trennung vom dynamischen und statischen Verhalten vorgenommen. Darüber hinaus wird eine Plausibilisierung und Kompensation von Ausreißern durchgeführt. Sensorfusion kann definiert werden als das Kombinieren und Zusammenführen von Messwerten von verschiedenen Sensoren zu einem Gesamtbild, unter der Berücksichtigung von Messunsicherheit und Messfehler, z. B. bei der Bestimmung der aktuellen Position aus mehreren Abstandsmessungen. Das Gesamtbild wird über Zustandsgrößen beschrieben. Der Zustand eines Systems wird durch vordefinierte Größen, wie z. B. die Position, Winkellage oder Geschwindigkeit definiert. Bei den meisten technischen Prozessen und Geräten ist es nicht möglich, die gewünschte Information über den Systemzustand präzise, vollständig, und robust mit nur einer einzigen Informationsquelle zu erfassen. Die Lösung zahlreicher Aufgaben setzt deshalb die Ausnutzung und Vereinigung unterschiedlicher Informationsquellen voraus.

Statistische Analyseverfahren - insbesondere die maschinellen Lernverfahren - eignen sich insbesondere zur Plausibilitätssteigerung sowie der Extraktion höherwertiger Information, d. h. dem automatisierten Erkennen fachtechnischer Sachverhalte. Die besondere Eigenschaft der maschinellen Lernverfahren liegt im Gegensatz zu den modellbasierten Analyseverfahren in ihrer selbstständigen Lernfähigkeit. So können diese Verfahren basierend auf einem Trainingsdatensatz selbstständig bestimmte Aufgaben

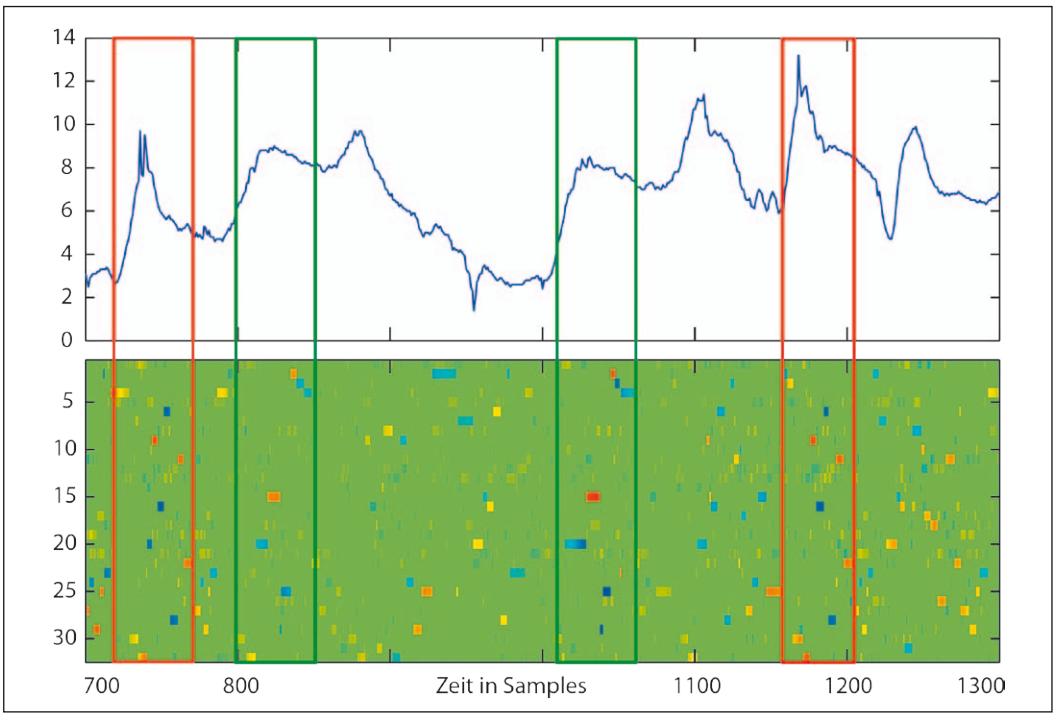

Bild 3: Extraktion von höherwertigen Information bzw. Mustern bei einem Fertigungsprozesses. Im oberen Bild sind die Prozessdaten dargestellt und im unteren Bild das Ergebnis der Klassifikationsalgorithmen.

erlernen, ohne explizit dafür programmiert zu werden - es ist kein physikalisches Modell erforderlich. Die Leistungsfähigkeit, und die Möglichkeit der Erfüllung der gestellten Aufgaben, hängen von der Struktur und Anordnung der Verfahren ab. Die Herausforderung bei der Entwicklung derartiger Algorithmen besteht darin, die Struktur für bestimmte Aufgabenstellungen zu finden und zu optimieren. Ist diese einmal für eine bestimmte Problemklasse bzw. Sensortyp definiert, können die Verfahren diese Struktur automatisch durchführen und dessen Parametrierung und somit Verhalten selbstständig erlernen.

- **Principal Component Analysis (PCA)**, Hauptkomponentenanalyse, ist ein Verfahren aus der multivariaten Statistik. Basierend auf einer mittels Trainingsdaten gelernten Linearkombination von Daten kann z. B. ein erwarteter Messwert und somit die Abweichung zum tatsächlichen Messwert berechnet werden.
- **Künstliche Neuronale Netze (KNN)** sind mathematische Beschreibungen welche versuchen, die Struktur und Informationsarchitektur eines Nervensystems von Tieren oder Menschen näherungsweise abzubilden. Die besondere Eigenschaft der künstlichen Neuronalen Netze ist ihre selbstständige Lernfähigkeit, d. h. basierend auf einen Trainingsdatensatz können diese selbstständig bestimmte Aufgaben erlernen ohne explizit dafür programmiert zu werden.
- **Self-Organizing Maps (SOM)** sind eine besondere Art von künstlichen neuronalen Netzen für das unüberwachte Lernen von Merkmalen bezogen auf abgegrenzte Datengruppen. Die gelernten Merkmale können zur Überprüfung der Plausibilität der Mess- und Erfassungsdaten, d. h. der Erkennung von Auffälligkeiten genutzt werden.
- **Generative Topographic Mapping (GTM)** kann als probabilistische Erweiterung von SOMs gesehen werden. Unsicherheiten in Mess- und Erfassungsdaten und dem gelernten Modell können systematisch berücksichtigt werden.

In der Praxis findet die Datenverarbeitung typischerweise in zwei Phasen statt: Der Lernphase, und der Erkennungsphase. In der Lernphase wird vollautomatisch ein Modell der Messdaten generiert. Hierbei werden Lernverfahren eingesetzt, um das Normalverhalten des Systems zu lernen. Dieses Normalverhalten wird im Verfahren so repräsentiert, dass Aspekte der zeitlichen und örtlichen Abhängigkeit kodiert werden. Sobald die Lernphase abgeschlossen ist, ist das Verfahren in der Erkennungsphase in der Lage, Abweichungen vom gelernten Modell in Echtzeit zu erkennen. Abweichungen des Modells werden dem Benutzer als Anomalie-Indikatoren zur Verfügung gestellt.

Anwendungsbeispiele

Der Hauptnutzen von Industrie-4.0-Anwendungen ist die Qualität der Produkte und die Effizienz der Produktion zu erhöhen. Eine intelligente Datenverarbeitung trägt an unterschiedlichsten Stellen der Wertschöpfungskette erheblich dazu bei dieses Ziel zu erreichen.Zur Qualitätssicherung können aus den Prozessdaten fehlerhafte Produkte automatisch detektiert und frühzeitig aus dem weiteren Produktionsprozess genommen werden. Dies reduziert einmal den Ressourcenbedarf und trägt zu einer höheren Qualität der Produkte bei. Spezielle Algorithmen zur Anomalie-Erkennung detektieren automatisch Veränderungen von einem Nominalzustand eines Prozesses mittels den aufgezeichneten Prozessdaten und deren Verarbeitung. Durch die Verwendung dieser Software kann eine permanente Überwachung der korrekten Funktionalität und das frühzeitige Erkennen von fehlerhaften Produkte sichergestellt werden. Weiter können Fehler und Verschleiß der eigentlichen Maschinen vor einem Ausfall erkannt und somit Wartungsintervalle optimal geplant werden. Damit werden die Prozesse zur Qualitätssicherung stetig verbessert und effizient gestaltet.

Werden im Rahmen der automatischen Prozessüberwachung fehlerhafte Produkte erkannt, so kann in einem zweiten Schritt eine (automatische) Fehlerdiagnose durchgeführt werden. Dies beinhaltet, dass ähnlich auftretende Veränderungen zu einem gemeinsamen Ereignis zusammengeführt werden, welches darüber hinaus in der Trainingsphase zusätzliches Expertenwissen beinhalten muss. In der Einsatzphase bestimmen spezielle Algorithmen zur Fehlerklassifikation automatisch die Art und die Schwere des Ereignisses und ermöglichen es, ggf. Handlungsanweisungen abzuleiten. Durch diesen Ansatz ist eine automatische Prozessplanung, und -steuerung möglich.

Fazit

In diesem Artikel wurden Chancen und Herausforderung der Datenverarbeitung in Industrie-4.0-Anwendungen dargestellt. Neben vielen Chancen besteht nach Erfahrung der Autoren die größte Herausforderung in dem unterschiedlichen Verständnis und Sprache der informationstechnisch geprägten und der produktionstechnisch geprägten Unternehmen. In den meisten Fällen ist es so, dass die Unternehmen aus dem Produktionsumfeld zwar die eigenen Produktionsanlagen bis ins letzte Detail kennen und auch eine sehr gute Vorstellung davon haben, welche zusätzliche Prozessinformation das Potential für weitere Effizienzsteigerung besitzen, aber nicht wissen wie sie aus der Vielzahl an Prozessdaten die gewünschten Information extrahieren können. Hier besteht großes Potential in der der engen Kooperation mit spezialisierten, informationstechnisch geprägten Anbietern, die verschiedenste Algorithmen entwickeln und softwaretechnisch umsetzen können.

■ *Knowtion UG*
www.knowtion.de