

ビッグデータを活用した歩留まり解析支援システム “歩留まり新聞”

Comprehensive Monitoring System for Yield Analysis Using Big Data in Semiconductor Manufacturing

中田 康太 NAKATA Kota 折原 良平 ORIHARA Ryohei

半導体製造における歩留まり解析では、半導体製品の品質検査結果と各工程の処理履歴から不良の原因を特定し、対策につなげることで生産性向上に寄与している。製造プロセスから得られるデータは大量かつ複雑であるため、人手による解析では作業時間が長くなる上、有用な情報を読み取ることが難しいという問題があった。

東芝グループは、機械学習とデータマイニングの手法を適用し、不良発生状況の可視化と不良の原因となる装置の推定を網羅的に行う、歩留まり解析支援システム“歩留まり新聞”を開発した。歩留まり新聞は、ウエハー上のどこに不良が発生したかを示す不良マップを提示するのに加え、不良マップをその特徴により自動で分類し、それぞれの不良原因となる装置候補を抽出して不良原因装置の候補を提示する。半導体工場の実データを用いた実験により、不良1件当たりの解析時間を従来の平均6時間から2時間に短縮できることを確認した。

Yield analysis contributes to improvements in the productivity of semiconductor manufacturing through identification of the causes of failures and provision of the appropriate countermeasures based on the results of product inspection and processing histories. However, the large volumes of complicated data collected in each process require a long processing time, making it difficult to manually extract useful information from raw data.

In order to resolve this issue, the Toshiba Group has developed a monitoring system for yield analysis in semiconductor manufacturing that is capable of comprehensively visualizing a summary of failure occurrences and inferring candidates for failure-causing machines by applying machine learning and data mining techniques to big data processing. This yield monitoring system automatically classifies distinctive types of wafer failures, thereby offering information on wafer failure maps and candidates for failure-causing machines. We have conducted simulation experiments using actual data from our semiconductor manufacturing factory and confirmed that this system reduces the average turnaround time of yield analysis from the conventional six hours to two hours.

1. まえがき

最新の半導体工場では、大量の製造データが収集され、生産性の向上に利用されている。生産性向上の取り組みの一つである歩留まり解析では、人が製品の品質検査結果と各工程の処理履歴から不良の原因を特定し、対策を行うことで製品の歩留まりを向上させている。大量のデータを用いることで、過去には不可能だった詳細な歩留まり解析が可能になり、大幅な生産性向上が期待されている。一方で、データは膨大かつ複雑になっており、その中から生産性向上につながる有用な情報を読み取ることは、ますます困難になっている。

図1は、歩留まり解析が持つ課題を表している。一般的に、大規模な半導体工場では数十から数百の種類の商品を生産する。半導体製品はウエハーと呼ばれる円盤上に製造

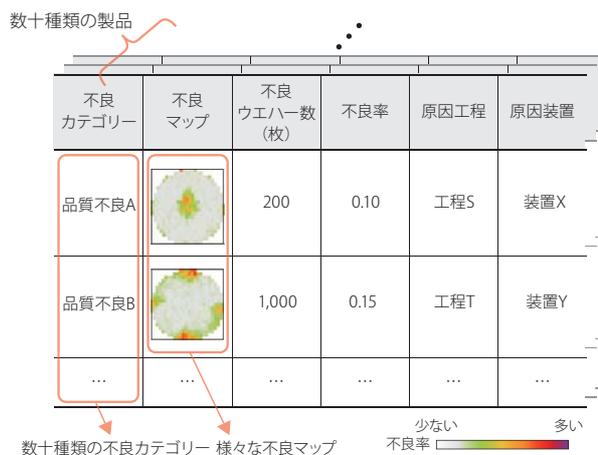


図1. 歩留まり解析の課題

多種多様な製品・不良カテゴリー・不良マップの組み合わせの中から、不良の原因となっている装置を特定する必要があり、作業コストが大きい。

Issues in comprehensive monitoring for yield analysis

し、それぞれの製品で品質検査を行い、ウエハーから切り出した各製品チップに対して、数十種類の検査項目で不良の有無を判定する。製品チップを不良と判定した場合には、その検査工程の名前を不良カテゴリーとして付与する。製造過程でトラブルが起こった場合には、ウエハー上のどこにあったチップが不良だったかを示す不良マップに特徴的な傾向が出現することが知られている。不良マップは製品によって多様であり、過去にない新たな不良マップが発生することもある。製品、不良カテゴリー、不良マップの組み合わせは膨大になるため、その中から重要な不良を素早く発見して原因の特定を行うことが、生産性向上への大きな課題の一つとなる。

この課題に対し、機械学習及びデータマイニングの技術を用いて、不良マップの発生状況の可視化と不良原因装置の推定を網羅的に行う歩留まり解析支援システム“歩留まり新聞”を開発した¹⁾。歩留まり新聞は、特徴的な不良マップを自動で分類し、それぞれの分類に対応する不良原因装置の候補を抽出する。また、解析が完了した不良マップを既知の不良としてシステムに学習させることで、同じ不良の再発を自動で監視できる。

ここでは、機械学習及びデータマイニング技術を活用して新規に開発した不良マップの自動分類、不良原因装置の推定、及び既知不良マップの再発監視と、これらを統合した歩留まり新聞について述べる。

2. 歩留まり解析の自動化技術

2.1 不良マップの自動分類

歩留まり新聞では、不良マップを自動分類するためにクラ

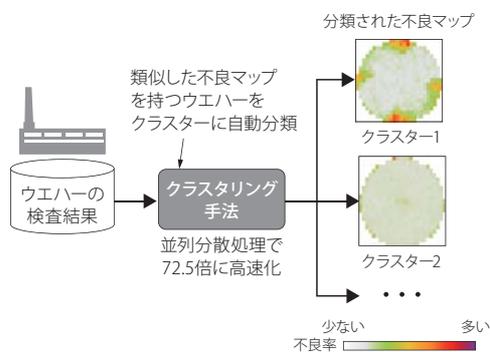


図2. クラスタリング手法を用いた不良マップの自動分類

並列分散クラスタリング手法により、類似した不良マップを自動的にクラスター（グループ）に分類する。

Automated classification of wafer failure maps by means of parallel distributed clustering

スタリング手法を導入した。クラスタリングはデータ間の距離に応じてデータをクラスターと呼ばれるグループに分割する手法で、機械学習のための人による教師信号を必要としないメリットもあり、様々な分野で利用されている。図2は、クラスタリング手法を用いた、不良マップの自動分類の手順を表している。ここでは、不良マップが類似したウエハーを分類して可視化することで、1枚1枚のウエハーを目視で確認することなく、不良マップと発生枚数を把握することができる。

大規模な半導体工場では、日々大量の製品が生産されているため、クラスタリングの精度だけでなく処理速度も重要な要素になる。歩留まり新聞では、全製品・全不良カテゴリーのクラスタリングをタイムリーに完了するために、並列分散処理が可能なクラスタリング手法 K-Means++を導入した²⁾。32並列の計算ユニット上でK-Means++を用いることで、従来のクラスタリングの処理速度を、最大で72.5倍に高速化した。これにより、全製品・全不良カテゴリーに対し、数時間で処理を完了でき、前日までの不良の発生状況を毎朝確認することが可能になった。

2.2 不良原因装置の推定

各不良マップに対応する原因装置を自動推定するために、パターンマイニング手法を導入した。パターンマイニング手法の代表的な応用としては、顧客の買い物履歴（トランザクション）から、同時に購入される商品のパターンを抽出する例がよく知られている。図3は、不良原因装置の自動推定に用いるデータとパターンマイニング手法の概要を表している。ここでは、ウエハーが通ってきた各工程の処理装置履歴をトランザクションとみなして、不良マップを持つウエハーに

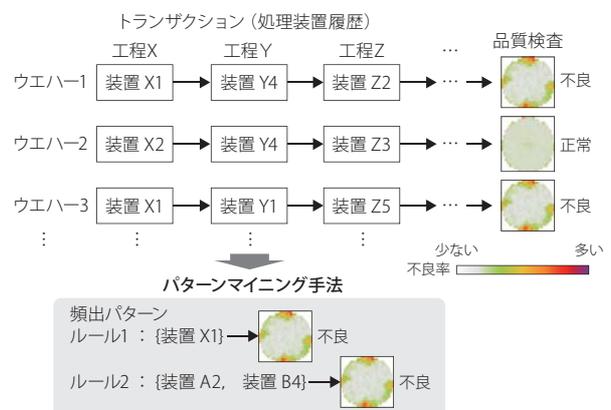


図3. パターンマイニング手法による不良原因装置の抽出

不良が発生するウエハーの処理履歴に頻出する装置パターンを抽出する。

Extraction of candidates for failure-causing machines by means of pattern extraction method

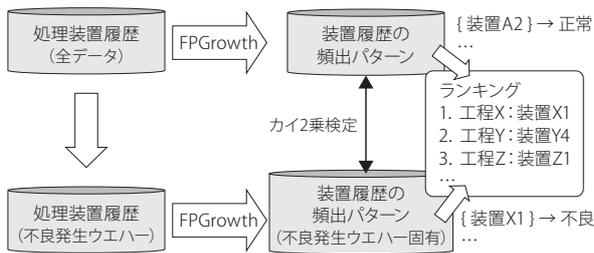


図4. 不良原因装置の推定フロー

正常ウエハーと不良ウエハーでそれぞれ頻出する装置群を抽出し、差分が大きい装置を不良原因装置の候補としてランキング表示する。

Flow of identification of failure-causing machines

固有の装置の頻出パターンを抽出することで、不良原因装置の推定を行う。

図4は、歩留まり新聞が不良原因装置を推定するフローを表している。全ウエハーの処理装置履歴と、不良マップが発生しているウエハーの処理装置履歴から、それぞれ装置履歴の頻出パターンを抽出する。ここではパターンマイニング手法としてFPGrowthを利用する⁽³⁾。FPGrowthは、大規模なデータから効率的に頻出パターンを抽出できるのが特長である。

次に、抽出された、それぞれの頻出パターンの中でカイ2乗検定を行い、不良が発生しているウエハーに特有の装置パターンを不良原因装置の候補としてランキング形式で表示する。この構成は、高速での分散処理も可能なため、全製品・全不良カテゴリー・全不良マップに対して不良原因装置を推定できる。

実際に発生した四つの不良事例について、この手法を適

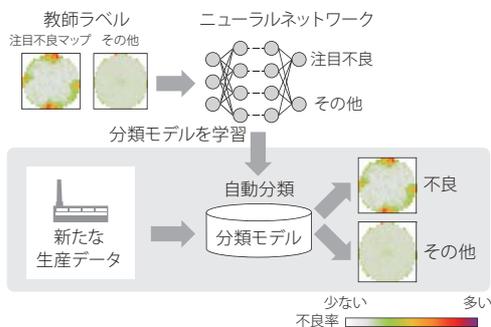


図5. 既知不良マップの再発監視

技術者が与えた教師ラベルを基に学習した分類モデルで、新たなウエハーを自動分類して再発を監視する。

Monitoring of recurrence of prior failures

用したところ、真の不良原因装置を3位以内に提示できることを確認した。現在は、人からのフィードバック情報を基に継続的に推定結果を評価し、推定精度の向上に取り組んでいる。

2.3 既知不良マップの再発監視

図5は、歩留まり新聞を用いた、既知不良マップの監視の概要を表している。ここでは、標準的な教師あり学習のアプローチで、既知の不良マップの再発監視を行う。ウエハーに対して注目する不良の発生有無を、人手で教師ラベルとして付与し、その教師ラベルを基に学習した分類モデルを用いて、新たなウエハーを分類して再発を検知する。

学習手法としてディープラーニング(深層学習)を用いる。ディープラーニング手法は、画像認識をはじめとして多くの応用分野で成功を収めている。不良マップは、シンプルな画像とみなせるので、ディープラーニング手法による高性能な自動分類が期待できる。一方で、ディープラーニング手法で高い性能を得るためには、ネットワーク構造を含む多くのパラメータを適切に設定する必要がある。実データを用いて実験を行った結果、特に性能に対して影響の大きかったネットワーク構造について、以下で説明する。

この実験では、半導体工場で得られた、2か月分のウエハーのデータを用いた。現実の再発監視を模擬するために、前半の1か月のデータを学習用、後半の1か月のデータを評価用とし、学習用のデータに注目する不良マップかどうかを教師ラベルとして与え、注目不良の発生ウエハーとそれ以外のウエハーの2値で分類を行う分類モデルを学習した。分類モデルにより評価用データを分類し、不良ウエハーに対するF値により学習の性能を評価した。ここで、F値とは、分類の正確さを表す適合率(不良と分類されたウエハーに対し、実際に不良であったウエハーの割合)と、分類の網羅性を表す再現率(実際の全不良ウエハーに対し、不良と分類されたウエハーの割合)の調和平均である。

表1は、ニューラルネットワークの構造と分類精度の関係を表している。サポートベクターマシン(性能が高いことが知られている学習手法の一つ)と、5層、7層、及び9層の

表1. ニューラルネットワークの実データによる分類精度の実験結果
Results of experiments on performance of deep neural networks using actual manufacturing data

手法	SVM	5層NN	7層NN	9層NN
精度 (%)	88.4	92.0	90.6	91.5

SVM: サポートベクターマシン NN: ニューラルネットワーク

ニューラルネットワークとを用いて学習を行った。ディープラーニング手法の性能が大きく上回っていることが分かる。

一般画像認識の分野では、数十層の構造を持ったネットワークにより、高い分類性能が得られることが知られているが、今回の実験では5層のニューラルネットワークが最も高い性能を示している。不良マップは、比較的シンプルな画像であるため、シンプルなネットワークで学習を行うことが適切であると考えられる。

3. 歩留まり新聞

図6は、2章で述べた技術を組み込んだ歩留まり新聞による、不良マップ分類と不良原因装置の推定結果の表示例を表している。不良マップや、不良が発生したウエハーの枚数、不良原因装置の候補のサマリーが確認できる。新たな不良が発生した場合には、不良原因装置の候補に対し、人

製品	不良 カテゴリ	不良 マップ	不良発生 トレンド	不良 ウエハー 数(枚)	不良原因装置 の候補
製品P	カテゴリ 1			300	1 工程 X: 装置 X1 2 工程 Y: 装置 Y4 3 工程 Z: 装置 Z9
製品Q	カテゴリ 2			200	1 工程 A: 装置 A5 2 工程 B: 装置 B2 3 工程 C: 装置 C4
製品R	カテゴリ 3			100	1 工程 D: 装置 D4 2 工程 E: 装置 E2 3 工程 F: 装置 F7

少ない 多い
 不良率

図6. 歩留まり新聞の表示例

不良マップの発見と不良原因装置の特定を、高速に実施できるようにした。
 Example of wafer failure map pattern monitoring/failure-causing machine identification display

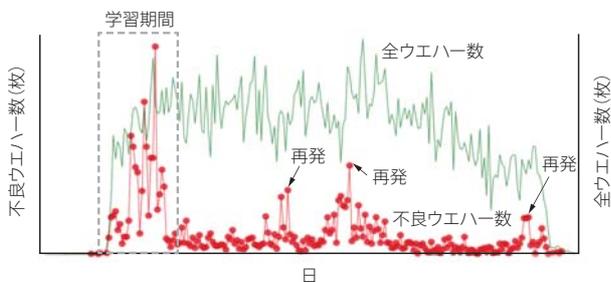


図7. 歩留まり新聞による既知不良マップの再発監視

長期間のトレンドを提示することで、容易に再発を発見することができる。
 Example of prior failure recurrence monitoring display

による詳細な調査を開始することが可能であり、歩留まり解析作業に掛かる時間を大幅に短縮できる。高速なクラスタリングとパターンマイニング手法により、全製品・全不良に対してサマリーの出力が可能になった。

図7は、ディープラーニング手法を用いて、既知不良マップの再発監視を表示した例である。グラフ中の点線で囲んだ期間にあるウエハーの情報を用いて学習を行った。緑線は1日に製造された全体のウエハー数を、赤点は注目する不良が起きていると判定されたウエハーの枚数を表している。この表示により、長期の不良発生トレンドの監視が容易になるので、不良の再発に迅速に対応できる。

4. あとがき

歩留まり新聞により、毎朝、前日までの全製品・全不良カテゴリーの不良マップや、不良発生トレンド、不良原因装置の候補、再発の情報を一覧で確認することができる。新しい不良マップと不良原因装置の候補を提示することで、不良1件当たりの解析時間を平均6時間から2時間に短縮した。

今後は、システム上に蓄積される人の知識を活用することで、更なる分析時間の短縮を目指す。また、半導体以外の製造データにも適用範囲を広げ、様々な分野での生産性向上に役立てていく。

文献

- (1) Nakata, K. et al. A Comprehensive Big-Data-Based Monitoring System for Yield Enhancement in Semiconductor Manufacturing. IEEE Transaction on Semiconductor Manufacturing. 2017, 30, 4, p.339-344.
- (2) Bahmani, B. et al. Scalable k-means++. Proceedings of the VLDB Endowment. 2012, 5, 7, p.622-633.
- (3) Han, J. et al. Mining frequent patterns without candidate generation. ACM SIGMOD record. 2000, 29, 2, p.1-12.



中田 康太 NAKATA Kota, Ph.D
 研究開発本部 研究開発センター アナリティクスAIラボラトリー
 博士(理学)
 人工知能学会会員
 Analytics AI Lab.



折原 良平 ORIHARA Ryohei, Ph.D
 研究開発本部 研究開発センター アナリティクスAIラボラトリー
 博士(工学)
 人工知能学会・情報処理学会・日本ソフトウェア科学会会員
 Analytics AI Lab.