

Aprendizagem de máquinas e correspondência de padrões em design físico

Bei Yu¹, David Z. Pan¹, Tetsuaki Matsunawa², e Xuan Zeng³

Departamento IECE, Universidade do Texas em Austin, Austin,

TX, EUA

²Centro de Pesquisa e Desenvolvimento de Semicondutores, Toshiba Corp., Kawasaki, Japão

³Laboratório de chaves estaduais do ASIC & Systems, Departamento de Microeletrônica,
Universidade de Fudan, China

{bei,dpan}@cerc.utexas.edu, tetsuaki.matsunawa@toshiba.co.jp, xzeng@fudan.edu.cn

ABSTRACT

A aprendizagem de máquinas (ML) e a correspondência de padrões (PM) são técnicas computacionais que podem derivar conhecimento de grandes dados, e fornecer previsão e correspondência. Desde que o projeto e a fabricação do nanômetro VLSI têm uma complexidade extremamente alta e dados gigantescos, tem havido uma re

centrado na aplicação e adaptação de aprendizagem de máquinas e técnicas de correspondência de pattern em projeto físico VLSI (incluindo verificação física), por exemplo, detecção de hotspot litográfico e projeto físico orientado por dados/padrões, já que ML e PM podem elevar o nível de abstração de simulações detalhadas baseadas em física e fornecer uma qualidade razoavelmente boa de resultado. Neste artigo, discutiremos técnicas-chave e resultados recentes da aprendizagem de máquinas e da correspondência de padrões, com suas aplicações em design físico.

I. INTRODUÇÃO

Como o tamanho das características dos nós de tecnologia de processo de semicondutores diminui ainda mais, a indústria é muito desafiada em termos de fabricação e design [1]. Por um lado, uma vez que a tecnologia litográfica principal ainda é limitada pela litografia de 193nm de comprimento de onda, o que-you-see no projeto não é necessariamente o que-you-get na fabricação. Mesmo com as regras de design complicado, várias técnicas de melhoria de resolução (RETs), e técnicas de litografia de padrões múltiplos (MPL) niques, ainda pode haver complexos hotspots de litografia dependentes do layout que podem causar aberturas, shorts e perda de rendimento. Portanto, como detectar e remover hotspots litográficos durante o projeto físico é fundamental para garantir um alto rendimento. Por outro lado, como os circuitos VLSI modernos têm bilhões de transistores/interconexões, é simplesmente impossível executar litografia detalhada e outras simulações baseadas em física na escala completa do chip durante a fase de projeto físico para lidar com hotspots litográficos

Recursos de
treinamento

Treinamento /
Calibração

Características
de teste

Teste /
Avaliação

potenciais e outras restrições de projeto.

O aprendizado de máquinas e técnicas de combinação de padrões proporcionam uma abstração razoavelmente boa e qualidade de resultado, o que as torna adequadas para realizar a detecção de pontos críticos litográficos, como demonstrado por muitos estudos recentes. Enquanto isso, os princípios de aprendizagem da máquina e de correspondência de padrões também têm sido usados para roteamento consciente da litografia e outras áreas de projeto físico, tais como a colocação de datapath e a otimização do relógio com sucesso. Neste artigo, apresentaremos algumas técnicas de aprendizagem e correspondência de padrões comumente usadas na máquina, e discutiremos

suas aplicações em projeto físico e verificação.

O restante do trabalho está organizado da seguinte forma. Na Seção II introduzimos alguns algoritmos básicos de aprendizagem de máquinas e correspondência de padrões. Na Seção III, discutimos as aplicações físicas de sinais. Na Seção IV discutiremos ainda algumas questões avançadas. A Seção V conclui este trabalho.

II. ALGORITMOS BÁSICOS

A. Técnicas de aprendizagem de máquinas

A aprendizagem de máquinas é uma disciplina da informática que lida com a construção e estudo de algoritmos que podem aprender com os dados [2]. As técnicas de aprendizagem de máquinas amplamente aplicadas no projeto físico são chamadas de aprendizagem supervisionada, onde os dados de treinamento contêm exemplos explícitos do que a saída correta deve ser para determinadas entradas. Um fluxo de aprendizagem

supervisionado clássico é ilustrado na Fig. 1 que consiste de treinamento/calibração e estágios de teste/avaliação. Na etapa de treinamento, dadas as características do treinamento de entrada, um conjunto de modelos de aprendizagem é construído. Na fase de teste, os modelos de aprendizagem construídos devem tomar pré-dições ou decisões, em vez de seguir apenas instruções gramaticais explícitas. Nesta subseção, introduzimos alguns modelos de aprendizagem típicos e populares na aprendizagem supervisionada.

A.1 Rede Neural Artificial (ANN)

A rede neural artificial (ANN) é inspirada pelo cérebro humano para estimar ou aproximar funções que podem depender de um grande número de entradas e os detalhes das funções são geralmente não

conhecido. ANN cria uma rede nivelada de neurônios, e cada neurônio i pode ser treinado para modelar uma função li :

$\times \rightarrow Y$ através do algoritmo de retropropagação [3]. Um ANN com mul \Rightarrow Learning Models \Rightarrow Os níveis da ponta podem assim modelar funções altamente não lineares. As características de entrada são apresentadas à ANN através da camada de entrada, que

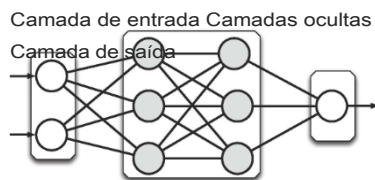


Fig. 2. Um exemplo de ANN de quatro camadas.

se comunica com uma ou mais camadas ocultas onde o processamento real é feito. Em seguida, as camadas ocultas são conectadas a uma camada de saída, que dimensiona sua entrada para a resposta desejada. Um exemplo de uma ANN de avanço de quatro camadas é mostrado na Fig. 2.

Em comparação com outras técnicas de aprendizagem de máquinas, a ANN geralmente tem uma boa resistência ao ruído. Entretanto, leva mais tempo no treinamento e na calibração para alcançar soluções ótimas ou próximas das ideais.

A.2 Máquina Vetorial de Apoio (SVM)

A máquina vetorial de suporte (SVM) é uma das técnicas de classificação e aprendizagem mais populares. Na SVM, os dados vetoriais são mapeados em um espaço de maior dimensão usando uma função kernel, e uma função de discriminação linear ótima no espaço ou um hiperplano ótimo que se encaixa nos dados de treinamento é construído [4]. O objetivo é maximizar a margem entre o hiperplano separador e os vetores de dados mais próximos de ambas as classes. Se os dados não forem separáveis, um SVM tipo C pode ser aplicado ao problema de classificação da seguinte forma [5].

$$\begin{aligned} \min_{\alpha} & \quad \alpha^T Q \alpha - e^T \alpha \\ \text{s.t.} & \quad z^T \alpha = 0 \\ & \quad Q_{ij} = z_i z_j K(X_i, X_j) \quad i, j = 1, \dots, n \\ & \quad \alpha_i \in [-C, C] \end{aligned} \quad (1)$$

onde $X_i \in \mathbb{R}^d$ são vetores de treinamento, $z \in \{1, -1\}$ é o vetor indicador e e é o vetor de todos uns, Q é uma matriz $n \times n$ positiva semidefinida, e o parâmetro C controla o trade-off entre permitir erros de treinamento e forçar margens rígidas de separação. Para cada elemento $Q_{ij} \in Q$, $Q_{ij} = z_i z_j K(X_i, X_j)$.

A função kernel K mapeia os dados para o espaço diferente para que um hiperplano possa ser usado para fazer a separação. Depois de resolver o problema de otimização em (1), todos os $\alpha_i \in \alpha$ podem ser calculados. Então, dado um novo vetor de teste t , a função de decisão é $\text{sgn}(\sum_{i=1}^n \alpha_i K(X_i, t))$, onde a função sgn é definida como

$$\text{sgn}(x) = \begin{cases} -1 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}$$

Em teoria, a SVM garante o ótimo global, mas é sensível ao ruído de dados.

Fig. 3. Fluxo de correspondência de padrões clássicos.

estes classificadores fracos. Apesar de haver muitos que impulsionam alguns ritmos, um processo de aprendizagem é comum para a maioria deles. No processo de aprendizagem dos classificadores fracos, o peso dos dados de treinamento é atualizado iterativamente, de modo a aumentar seu peso para os dados de classificação errada e diminuir seu peso para os dados corretamente classificados. Um classificador final forte consiste em classificadores fracos que são aprendidos com dados de treinamento de diferentes pesos.

Embora o Boosting tenha muita flexibilidade no projeto do algoritmo de aprendizagem, o classificador fraco deve ser um classificador que pode ser treinado com dados ponderados. Os amplamente utilizados algoritmos para Boosting e o classificador fraco são AdaBoost [6].

Um classificador de árvore de decisão [7]. No AdaBoost, o classificador fraco Y_m é aprendido com o peso de dados $D_m(X_i)$, onde X são os vetores de treinamento $X_i \in \mathbb{R}^d$, $i = 1, \dots, n$, e m é o número

de classificadores fracos $m = 1, \dots, M$. Em seguida, a ponderação combinada eficiente é avaliada como $13m - E_m/(1 - E_m)$, onde E é a relação de classificação errada para vetores de treinamento ponderados X_i . O peso dos dados do próximo classificador fraco Y_{m+1} é reponderado como $D_{m+1}(X_i) = 13m D_m(x_i)$ quando os dados X_i são erroneamente classificados pelo classificador anterior Y_m e $D_{m+1}(X_i) = D_m(X_i)$ caso contrário. Finalmente, $D_{m+1}(X_i)$ é normalizado, e o final

O classificador forte $Y(x)$ é formado como resultado de Y_m ponderado por $\alpha_m = \ln(1/13m)$.

$$Y(x) = \text{sgn} \left[\sum_{m=1}^M \alpha_m Y_m(x) \right] \quad (2)$$

Geralmente, o Boosting é sensível aos dados, incluindo muito ruído e aberrações. Entretanto, pode ser superior em relação à questão do excesso de ajuste, em comparação com outras classificações algo rítmico.

B. Técnicas de correspondência de padrões

As técnicas de aprendizagem da máquina são para atribuir um padrão desconhecido a uma das aulas possíveis, enquanto o padrão combina

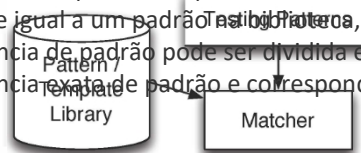
técnicas nesta subseção são de um pouco diferente na natureza. Um padrão típico de fluxo correspondente é ilustrado na Fig. 3,

onde um conjunto de padrões ou modelos está disponível na biblioteca. Difícil de aprender com a máquina, geralmente não há um trem

A.3 Boosting

O Boosting é outro método de aprendizado em conjunto para o problema de sificação de classes no aprendizado supervisionado. O conceito básico é que um conjunto de classificadores fracos é usado como um meio para criar um único classificador forte. Um classificador fraco é acusado de ser um classificador que tem um desempenho ligeiramente melhor do que a adivinhação aleatória. Um classificador forte é gerado por conjunção com

fase de construção da biblioteca. Dado um novo padrão de teste, precisamos decidir qual padrão na biblioteca combina melhor com o novo padrão. Dependendo se o padrão de teste é exatamente igual a um padrão da biblioteca, a correspondência de padrão pode ser dividida em correspondência exata de padrão e correspondência de padrão difuso.



B.1 Correspondência exata de padrões

Na correspondência exata do padrão, um padrão de teste é correspondido se e somente se houver um padrão exatamente igual na biblioteca. Se um conjunto

de características pode ser extraído de um padrão, a correspondência exata é o bem conhecido problema de correspondência de cordas [8]. Embora a correspondência exata do padrão seja de alto desempenho na detecção de padrões pré-determinados na biblioteca, falta-lhe a capacidade de identificar padrões nunca vistos.

B.2 Correspondência de Padrões Fuzzy

Em algumas aplicações, geralmente é difícil identificar um padrão que corresponda exatamente a um determinado padrão de teste. Em vez de re-portar "inigualável", em um problema de correspondência de padrão difuso, o padrão na biblioteca que melhor corresponda ao padrão de teste seria identificado. A correspondência de padrão difuso tem sido amplamente aplicada no reconhecimento de fala, detecção de rosto e processamento de dados [9,10]. O passo crítico neste problema é definir um custo para medir a "similaridade" entre dois padrões.

III. AÇÕES APLICÁVEIS EM FÍSICO DESENHO SICAL

Nesta seção mostraremos algumas aplicações chave da machine learning e da correspondência de padrões no projeto físico (incluindo a verificação física).

A. Detecção de pontos de acesso litográfico

A.1 Abordagem de aprendizagem da máquina

Nos estágios de projeto físico e verificação, o problema de detecção de hotspots é localizar hotspots em um determinado layout com tempo de retorno rápido. Simulação convencional de litografia [11,12] obtém imagens de padrões de banho usando modelos litográficos complicados. Embora seja precisa, a simulação litográfica de chip completo é computacional e cara, e, portanto, não pode fornecer um retorno rápido para guiar os estágios iniciais do projeto físico. A detecção de pontos críticos desempenha um papel essencial para preencher a grande lacuna entre a modelagem e a ferramenta de desenho físico consciente do processo.

Houve muitos trabalhos de detecção de hotspots baseados na aprendizagem de máquinas. Técnicas de aprendizagem de máquinas constroem um modelo de regressão baseado em um conjunto de dados de treinamento. Este método pode naturalmente identificar hotspots previamente desconhecidos. Entretanto, ele pode gerar falsos alarmes, que não são hotspots reais. Como melhorar a precisão de detecção é o principal desafio ao adotar técnicas de aprendizagem de máquinas.

Muitas abordagens recentes utilizam técnicas SVM e ANN para construir o núcleo de detecção de hotspots. Em [13], uma transformação em 2-D e extração de histograma é realizada em imagens de layout baseadas em pixels, que são usadas para construir a detecção de hotspots baseada em SVM. [14] apresenta um fluxo de detecção baseado no julgamento da rede neural, onde padrões de hotspots 2-D são usados diretamente para treinar um modelo ANN. [15] propõe um método híbrido e multinível adotando tanto SVM quanto ANN para melhorar o desempenho. Em [16,17], o SVM é empregado através da extração e classificação de métricas baseadas em densidade de layout. Em [18], as

metodologias baseadas na aprendizagem da máquina são estendidas ao hotspot de detecção de auto-montagem dirigida (DSA). O novo modelo DSA com correspondência pontual e características de distância de segmento são propostos para um aprendizado robusto.

Em [19-21], a análise de componentes principais (PCA) é aplicada para extração de características e redução de dados. Combinando a PCA

com SVM pode ajudar a melhorar a precisão de detecção de forma significativa. Muito recentemente,[22] propõe um modelo de classificação baseado no Boosting. Através de um layout simplificado, a utilização do algoritmo de aprendizagem fracamente não-linear é capaz de detectar pontos críticos com precisão com baixo alarme falso.

A.2 Abordagem de correspondência de padrões

Os métodos baseados na correspondência de padrões também são amplamente aplicados na detecção de hotspots. Um gráfico de layout é proposto em [23] para refletir a variação do CD relacionado ao padrão. O gráfico resultante pode ser usado para encontrar hotspots incluindo características fechadas, características em forma de L e padrões complexos. O conceito de padrão de alcance [24] é proposto para incorporar especificações dependentes do processo, e é aprimorado em [25] para representar novos tipos de hotspots. Um padrão de alcance é um layout bidimensional de retângulos com especificações adicionais codificadas por cordas. Cada padrão de alcance está associado a um mecanismo de pontuação para refletir o problema das regiões árticas de acordo com o impacto do rendimento. Os padrões de pontos de acesso são armazenados em uma biblioteca pré-definida e o processo de detecção realiza a correspondência de cordas para encontrar os pontos de acesso. Esta abordagem é precisa, mas a construção de padrões de alcance depende de uma matriz de layout baseada em grade, e pode ser demorada quando o número de grades é grande. Yu et al. [26] propõem uma detecção de hotspots baseada na RDC extraíndo características topológicas críticas e modelando-as como regras de projeto. Portanto, a detecção de hotspots pode ser vista como um processo de verificação de regras através de um DRC em gine.

Em [27], é proposto um esquema de classificação de hotspots baseado na correspondência de padrões. Os

hotspots são classificados em clusters por métodos de mineração de dados. O hotspot representativo em cada cluster é então identificado e armazenado em uma biblioteca de hotspots para futura detecção de hotspots. A abordagem de classificação de hotspots em [27] se baseia em uma métrica de distância de diferentes amostras de padrões, que é definida como uma integral ponderada sobre a área onde um par de padrões de hotspots difere (XOR de padrões). É sensível para as pequenas variações ou deslocamentos das formas. Em [28], uma métrica baseada no Espaço Tangente Melhorado (ITS) é proposta para a classificação dos hotspots. É uma extensão dos métodos de espaço tangente bem desenvolvidos [29,30] na comunidade de visão por computador. A métrica ITS define uma métrica de distância de um par de polígonos, que é a norma L_2 da diferença das funções de giro correspondentes dos polígonos [29] [30]. O funcionamento de giro do polígono mede o ângulo da tangente no sentido anti-horário em função do comprimento normalizado do arco, medido a partir de algum ponto de referência do polígono. O ITS baseado no $ric\ ric$ é fácil de calcular e é tolerante com pequenas variações ou deslocamentos das formas. Com a métrica baseada em ITS, a classificação do ponto de acesso pode alcançar maior precisão.

A.3 Aprendizagem da Máquina Híbrida e Correspondência de Padrões

Alguns trabalhos combinaram tanto a combinação de padrões quanto os métodos de aprendizagem de máquinas. Recentemente, é proposta uma correspondência difusa com alguma técnica de aprendizagem em [31] que pode afinar dinamicamente regiões difusas apropriadas em torno de pontos quentes conhecidos no espaço multi-dimensional. A Fig. 4 mostra um exemplo com padrões conhecidos de pontos quentes e pontos não quentes em um espaço bidimensional. Um método de aprendizagem de máquina dividiria o espaço em

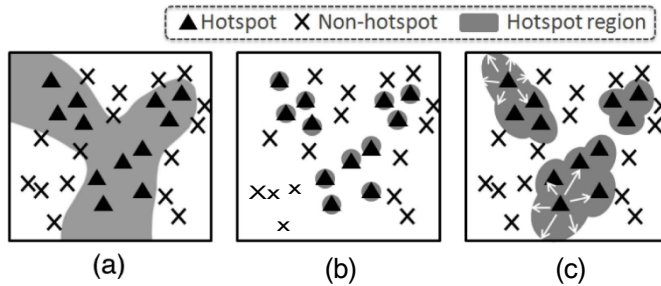
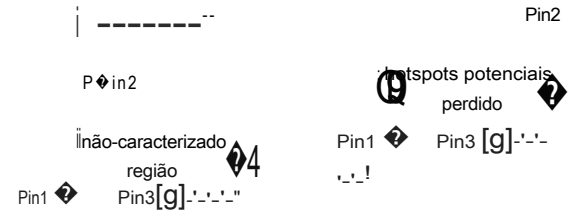


Fig. 4. Um exemplo em espaço 2D de decisão de região de hotspot. (a) Aprendizagem da máquina; (b) Combinação de padrões; (c) Modelo de correspondência Fuzzy. [31]



de pós-roteamento, por exemplo, [33]. Em [34] [35], o verificador de regras de projeto é integrado ao mecanismo de roteamento no estágio de pós-roteamento para identificar e corrigir hotspots. Entretanto, a fixação de pontos de acesso na fase de pós-roteamento tem flexibilidade limitada, já que apenas o rip-up é limitado.

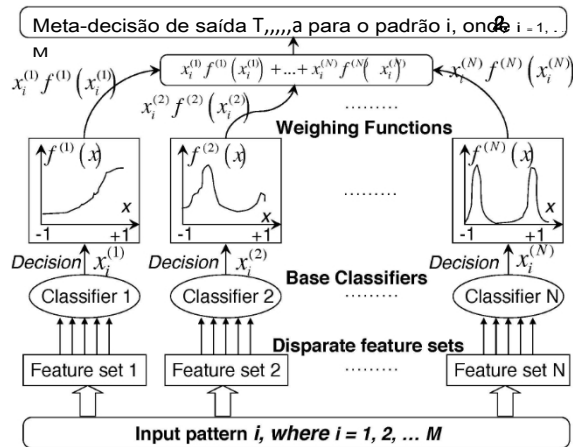


Fig. 5. Construção de metaclassificadores através de uma combinação de classificadores de base dispares [32].

duas regiões de hotspots e não hotspots como mostrado na Fig. 4(a), enquanto que uma abordagem convencional de correspondência de padrões conviria estruturar um padrão individual para corresponder a cada hotspot conhecido como mostrado em (b). O modelo de correspondência fuzzy na Fig. 4(c) inclui grupos de hotspots, onde a região fuzzy de cada grupo irá crescer iterativamente para proporcionar uma melhor precisão de detecção. Em [32], as amostras de dados são fornecidas primeiramente a um combinador de padrões, depois são usados classificadores de aprendizagem de máquina para examinar os pontos não detectados deixados pelo combinador de padrões. Motivado pelo fato de diferentes classificadores de pontos de acesso terem objetivos e pontos fortes diferentes[32], propõe ainda um metaclassificador unificado que permite que vários classificadores trabalhem em conjunto. A figura 5 ilustra o fluxo de construção da metaclassificação, que é composta de classificadores de base multiplataforma e funções de ponderação. Para cada padrão de layout, certas características do hotspot são extraídas e depois introduzidas em cada classificador de base, que calcula a decisão de previsão e gera um peso com base nas funções de ponderação. A meta-decisão final é baseada na soma pesada dos classificadores de base.

B. Litografia Roteiro Amigável

A mitigação do hotspot litográfico pode ser realizada na fase

Fig. 6. O desafio de detecção do hotspot na etapa de roteamento detalhado [36].

e redirecionamento pode ser realizado. Com previsões eficientes de hotspots, será interessante integrar o hotspot de técnica litográfica junto com o roteamento.

Um desafio do roteamento consciente da litografia é que os hotspots são difíceis de serem detectados antes que um caminho real de roteamento seja obtido. Fig. 6(a) dá uma região de layout com bloqueios de metal e pinos não roteados Pin1-Pin4. Como algumas redes ainda não são roteadas, existe uma região não caracterizada onde nenhum hotspot seria identificado por métodos gerais de detecção de hotspots. Consequentemente, os hotspots potenciais podem ser causados pela rota Pin1-Pin2, como na Fig. 6(b). [36] propõe um roteamento detalhado de litografia amigável baseado em um modelo pré-construído de previsão de hotspots e um modelo de previsão de rotas. O modelo de detecção de hotspot é treinado para avaliar a possibilidade de impressão de padrões com base em um conjunto de dados pós-RET. Para superar a questão de regiões não caracterizadas, o modelo de previsão de rotas é estabelecido usando as seguintes etapas: (1) explorar as possíveis soluções de roteamento dados os recursos de roteamento disponíveis; (2) realizar simulação litográfica precisa para os possíveis resultados de layout; (3) identificar as rotas preferenciais de acordo com os resultados de hotspots e congestionamentos de roteamento. Como os dados que precisam ser processados para a construção do modelo de

previsão de rotas são enormes, um classificador ANN é construído para guiar o mecanismo de roteamento.

C. Colocação do datapath

A principal objetivo típica da colocação é minimizar o comprimento total do fio de meio-perímetro (HPWL), que é um bom indicador de qualidade de colocação para projetos lógicos aleatórios. Entretanto, em projetos lógicos de datapath, geralmente as células são caracterizadas por um alto grau de paralelismo bit-wise que a colocação convencional tem mostrado ser sub-ótima [37]. Em projetos onde há muitos datapaths embutidos, extraíndo os datapaths e posicionando-os com lógica aleatória de forma apropriada tem o potencial de melhorar significativamente a StWL geral [38-40].

A figura 7 mostra um exemplo de brinquedo onde os colocadores modernos não são capazes de lidar eficazmente com os dados [38]. Fig. 7(a) mostra o circuito do datapath, onde os pinos de entrada/saída são fixos e há três bit-stacks correspondentes às células: {2,3,4,5}, {6,7,8,9}, e {10,11,12,13}. Fig. 7(b) mostra a solução de colocação PADE, onde cada bit-stack é firmemente embalado e alinhado produzindo uma solução StWL de 524. Fig. 7(c) exibe a solução de colocação a partir de Fast-Place3 [41] onde a pilha de bits não está cuidadosamente alinhada produzindo StWL de 612.

Um novo fluxo de colocação com extração automática de dados é proposto em [38], que avalia e classifica todos os caminhos de dados importantes e de primeira ordem, e os otimiza juntamente com a extração de dados gerais.

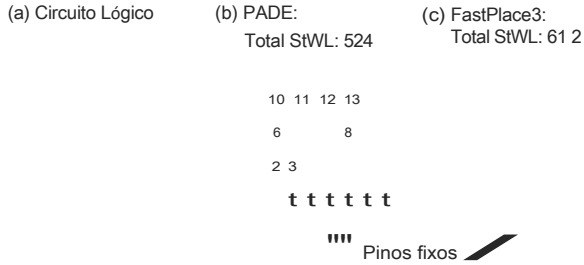


Fig. 7. Exemplo de posicionamento orientado por dados mostrando uma melhoria de 14% no comprimento do fio de steiner em comparação com o posicionamento convencional [38].

lingüeta. Em segundo lugar, um grande conjunto de modelos iniciais de posicionamento da lingüeta é gerado, e o número do modelo pode ser reduzido significativamente usando um conjunto teórico para remover a redundância. Por último, uma técnica de aprendizagem da máquina chamada indução de árvore de decisão é desenvolvida usando métrica de similaridade para selecionar rapidamente o gabarito correto durante a automação do projeto.

Fig. 8. Desenho Multi-GHz mostrando trincos agrupados, onde células vermelhas são trincos e células roxas são tampas de relógio locais (LCB) [45].

colocação com fio de comprimento de arame. Na fase de treinamento, para classificar e avaliar os padrões de datapath na lista inicial da rede, técnicas SVM e ANN são combinadas para construir modelos compactos e eficientes no tempo de execução. No modelo SVM, uma técnica tolerante a erros é combinada com uma seleção especial de conjunto de trabalho. ANN funciona através da configuração de redes complexas de neurônios para alcançar uma estrutura de dados de alta dimensão, semelhante a um diagrama de decisão, dadas as amostras de treinamento e as dicas de decisão. O objetivo da otimização tanto para SVM quanto para ANN é maximizar a precisão da avaliação dos padrões de dados e não-dados. Na fase de teste, os modelos de aprendizagem de dados serão aplicados diretamente para testar se um novo padrão de projeto desconhecido é um padrão de dados. Um padrão é avaliado para ser datapath se e somente se ambas as pontuações de avaliação SVM e ANN estiverem acima de certos limites.

D. Otimização do Relógio

O projeto da rede de relógios para microprocessadores de alto desempenho é um dos problemas mais desafiadores no projeto da VLSI. Para limitar a inclinação e o consumo de energia do estágio de nível inferior, as travas são frequentemente agrupadas e colocadas ao lado de um buffer de relógio local comum (LCB) de forma estruturada [42-44]. Como mostrado na Fig. 8, pequenos grupos de trincos são firmemente agrupados ao redor dos LCBs para reduzir significativamente o comprimento total da árvore do relógio local. [45] propõe uma solução escalável de aprendizagem de máquina para o problema de otimização da lingüeta. Primeiro, para cada aglomerado de lingüetas é proposto um algoritmo genético para procurar a solução do local da

(a)

(b)

Fig. 9. Recurso de layout baseado na fragmentação, onde o layout é cortado em fragmentos [15].



A. *Aprendizagem de máquinas ou correspondência de padrões?*

A aprendizagem da máquina, especialmente a aprendizagem supervisionada, consiste em etapas de treinamento e testes. Na fase de treinamento, dado o padrão de entrada, um conjunto de modelos de aprendizagem é construído. Em seguida, os modelos de aprendizagem construídos são aplicados na fase de teste para fazer previsões ou tomar decisões. Os métodos de análise baseados na aprendizagem de máquinas podem ser aplicados a classificações complexas não lineares, ou a aplicações onde as funções de classificação são desconhecidas. Entretanto, para manter um desempenho estável, a aprendizagem de máquinas requer um ajuste adicional de parâmetros para diferentes aplicações.

O método baseado na correspondência de padrões define e armazena padrões pré-caracterizados na biblioteca. Dado um padrão de teste, o padrão da biblioteca correspondente é identificado. O padrão de correspondência de padrões difusos pode ser aplicado para encontrar um padrão de biblioteca muito semelhante ao padrão de entrada. O método baseado na correspondência de padrões é muito rápido e preciso para detectar padrões conhecidos. Mas mesmo a correspondência de padrão felpudo é aplicada, pode faltar-lhe a capacidade de identificar padrões nunca vistos.

De modo geral, a aprendizagem de máquinas pode obter um valor maior por formalidade, especialmente quando se prevêem dados não vistos. Mas a machine learning pode levar mais tempo em treinamento para atingir um alto desempenho. Se os padrões da biblioteca forem suficientes e o tempo de execução for uma preocupação, o método baseado na correspondência de padrões é uma boa opção. As ferramentas de correspondência de padrões já foram integradas na indústria [46-49]. Por outro lado, a correspondência de padrões é mais sensível a mudanças de processo, pois um novo conjunto de padrões terá que ser construído após o ajuste do processo. O aprendizado da máquina pode ser mais roto sob mudanças de processo, e pode lidar com padrões invisíveis de forma mais natural. Serão desejáveis abordagens híbridas da aprendizagem de máquinas e da correspondência de padrões.

B. *Extração de recursos*

Tanto na aprendizagem de máquinas quanto nas técnicas de combinação de padrões, precisamos analisar os padrões. Cada padrão é extraído em uma característica específica, que é uma propriedade heurística individual mensurável que descreve o padrão. A extração de características é uma tarefa para representar os padrões usando informações simples, mas abrangentes, de modo que todas as características selecionadas possam ser distinguidas. Portanto, a extração de características é uma das etapas mais críticas tanto na aprendizagem da máquina quanto na correspondência de padrões. Nesta subseção, introduzimos e analisamos algumas técnicas clássicas de tração ex característica no projeto físico e na verificação física.

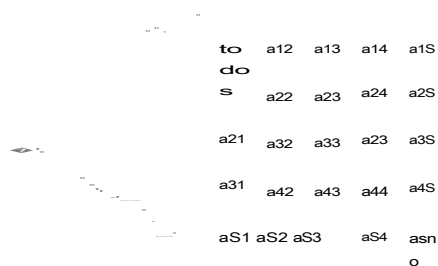


Fig. 10. Recurso de layout baseado em densidade. Característica Vector é representada como:

$X = \{\text{tudo}, a_{12}, \dots, a_{54}, a_{55}\} [22]$.

Fig. 12 compara o desempenho de diferentes características de tração, Podemos ver na Fig. 12 (a) que as características de layout baseadas na fragmentação estão separadas em vários clusters, e as características de hotspot estão misturadas com as características de não-hotspot. A partir daí, se não for cuidadosamente projetada, esta característica pode ser demasiado complicada para um único modelo de classificação para separar hotspots e não hotspots com alta precisão. A partir das Figuras 12 (b) e (c) podemos ver que, para este caso, a característica baseada em densidade e a característica baseada em HLAC podem definir facilmente os limites de decisão.

Na aplicação de detecção de hotspots, [22] propõe um método para

avaliar quantitativamente as características extraídas. O desvio médio das características do hotspot para as características que não são hotspot, z , é definido como segue. em consideração os propulsores de deslocamento. Em cada máscara, o preto representa os pixels envolvidos na multiplicação e o branco representa "não importa". As máscaras são usadas para varrer toda a imagem, culminando com o valor x em cada local e somando tudo. Os valores totais para todas as máscaras são concatenados a partir de um vetor como características HLAC.

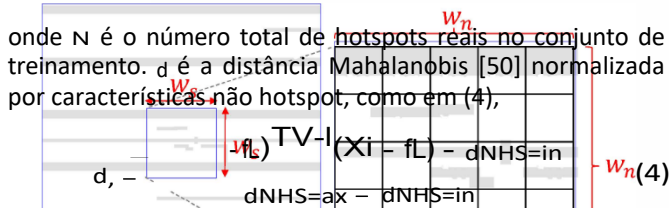
Fig. 11. Recurso de layout baseado em HLAC.

Um esquema de extração de layout baseado em fragmentação é apresentado em [15], onde o layout é cortado em fragmentos. Para cada fragment F , um raio r efetivo é definido para cobrir os fragmentos de perfuração que precisam ser considerados na caracterização do contexto de F , como mostrado na Fig. 9. Uma caracterização completa de F inclui a característica geométrica dos fragmentos dentro de r , incluindo formas de padrões, a distância entre os fragmentos, informações de cantos (convexos ou côncavos), e assim por diante,

As características de layout baseadas em densidade são apresentadas em [16,22], onde um padrão de layout é representado como um vetor de valores de densidade de layout de sua área ao redor. Dado um clipe de layout com grades pré-definidas, o método calcula a densidade de layout covarde em cada grade. Uma lista ordenada de valores de densidade forma então o vetor final que representa o padrão de layout correspondente. A Fig. 10 ilustra um exemplo de característica de layout baseado na densidade. O objetivo desta característica de layout não é identificar as características geométricas que podem degradar a capacidade de impressão de um padrão. Em vez disso, ela visa fornecer uma representação compacta dos padrões de layout para permitir a medição das semelhanças dos padrões.

Em [21] foi proposto um recurso de autocorrelação local de ordem superior (HLAC) baseado em [21]. Como ilustrado na Fig. 11, no layout baseado em HLAC, uma configuração de 3×3 pixels forma 25 padrões de máscaras locais que levam

(3)



De modo geral, $Z < 1$ indica que é difícil separar os hotspots dos não hotspots linearmente, já que a maioria dos hotspots está dentro do espaço de características não hotspot. Em contraste, $Z > 1$ mostra características amigáveis da separação linear devido à sua desatenção das características dos hotspots em relação às não hotspots, em outras palavras, para definir uma característica de layout apropriada, Z é de preferência maior que 1, mas não muito grande [22].

C. Superar a adaptação na aprendizagem de máquinas



291

O excesso de adaptação é o fenômeno em que a adequação dos moldes de treinamento não indica mais que os modelos de aprendizagem podem funcionar bem para os padrões de teste [51]. Em geral, o sobreajuste acontece quando os modelos de aprendizagem ou as características extraídas são mais complexos do que o necessário para representar o problema. Em outras palavras, se o modelo de aprendizagem utiliza graus adicionais de liberdade para encaixar o ruído nos padrões de treinamento, podemos observar uma má adaptação nos padrões de teste. Há várias razões pelas quais o excesso de adaptação acontece, ou seja, o ruído nos dados, a falta de amostras representativas, características muito complexas, e modelos de aprendizagem muito complexos [52].

A Fig. 13 mostra um exemplo de sobreajuste em um caso de teste de detecção de hotspot. A característica de layout baseado em densidade e os modelos de aprendizagem AdaBoost são utilizados em uma estrutura de detecção de hotspots. Para investigar o efeito do sobreajuste, diferentes níveis de complexidade podem ser aplicados. Quanto maior o valor w_n , mais complexo é extraído o recurso. Podemos ver que quando a complexidade da característica é muito baixa, tanto a precisão do treinamento quanto a precisão do teste são ruins, o que é conhecido como subequipamento do modelo. Como sempre, quando a complexidade da característica se torna muito grande, embora a precisão do treinamento ainda seja alta, a precisão do teste começa a diminuir.

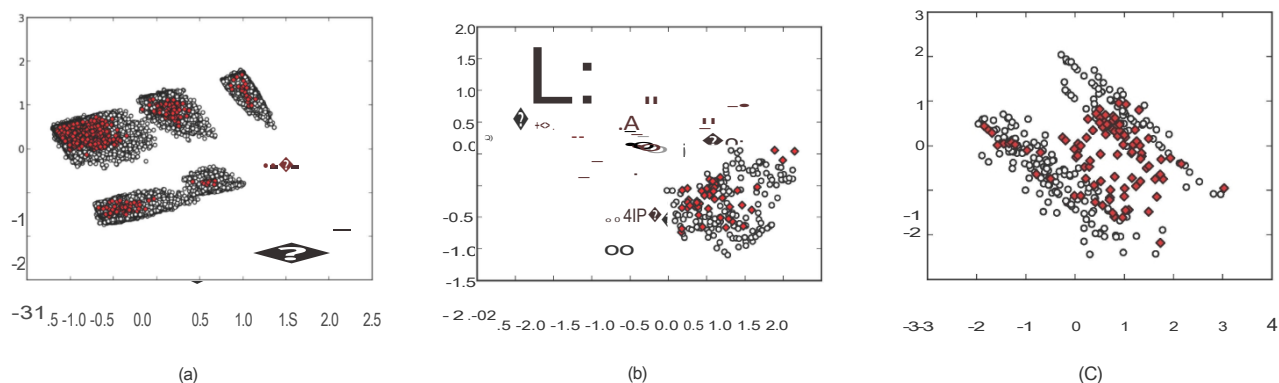


Fig. 12. Comparação de diferentes métodos de extração de características, onde os pontos vermelhos são características de hotspot e os pontos brancos são características de não hotspot: (a) Característica baseada em fragmentação; (b) Característica baseada em densidade; (c) Característica baseada em HLAC.

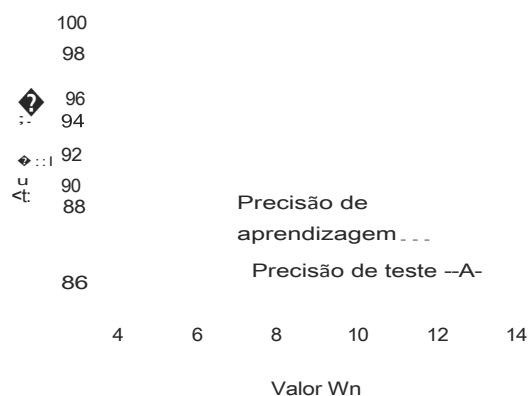


Fig. 13. Um exemplo de sobreajuste que uma vez a complexidade da característica (W_n) a precisão dos testes começa a diminuir.

Várias técnicas podem ser aplicadas para superar o problema do excesso de equipamento. A primeira é chamada de regularização, que em condições adicionais induz a restrições ou funções objetivas nos modelos de aprendizagem [53]. Por exemplo, a função objetiva em (1) pode ser reformulada como se segue.

(5)

onde A é um parâmetro para controlar a quantidade de regularização. O termo de penalidade $a.Ta$ impõe um trade-off entre fazer um treinamento mais preciso e um modelo de aprendizagem mais simples. Outro A técnica de sobreajuste é chamada de validação cruzada [54], onde os dados de treinamento são divididos em conjunto de treinamento e conjunto de validação. O conjunto de validação pode ser usado para estimar os dados de teste por formalidade. Além disso, o conjunto de validação também pode ser aplicado para selecionar modelos de aprendizagem apropriados.

V. CONCLUSÃO

Neste trabalho, pesquisamos algumas técnicas de aprendizagem e correspondência de padrões comumente utilizadas na ma chine. Também introduzimos suas aplicações no projeto físico e verificação, por exemplo, detecção de

As técnicas têm ganhado cada vez mais atenção para proporcionar uma abstração razoavelmente boa e qualidade de resultado. Esperamos que este trabalho estimule mais estudos sobre a aprendizagem de máquinas específicas de aplicação e técnicas de combinação de padrões em projeto físico e CAD VLSI.

AGRADECIMENTOS

hotspot litográfico, colocação de datapath, e otimização de relógio. Além disso, discutimos algumas questões avançadas, incluindo a extração de características e o problema de sobreajuste. Desde que os circuitos VLSI modernos têm bilhões de transistores/interconexões, aprendizagem de máquinas e combinação de padrões

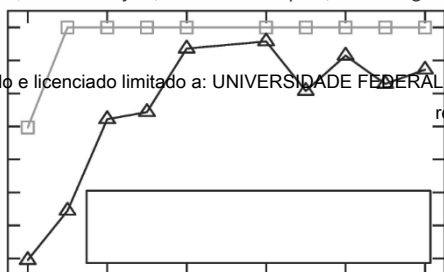
Este trabalho é apoiado em parte pela National Science Foundation (NSF), Semiconductor Research Corporation (SRC), Na tional Natural Science Foundation of China (NSFC) projeto de pesquisa 61125401, 61376040 e 61228401, Shanghai Science and Technology Committee projeto 13XDI401100, e Toshiba.

REFERÊNCIAS

- [1] D. Z. Pan, B. Yu, e J.-R. Gao, "Design para fabricação com nanolitografia emergente", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. 32, no. 10, pp. 1453-1472, 2013.
- [2] M. Mohri, A. Rostamizadeh, e A. Talwalkar, *Fundações da Aprendizagem de Máquinas*. MIT press, 2012.
- [3] M. T. Hagan, H. B. Demuth, e M. H. Beale, *Neural Network Design*. Pws Boston, 1996, vol. I.
- [4] B. E. Boser, I. M. Guyon, e V. N. Vapnik, "Um algoritmo de

treinamento para classificadores de margem ideal", em *Conference on Learning Theory*. ACM, 1992, pp. 144-152.

- [5] C. Cortes e V. Vapnik, "Redes de apoio-vetor", *Machine Learning*, vol. 20, não. 3, pp. 273-297, 1995.
- [6] Y. Freund e R. E. Schapire, "A decision-theoretic generalization of on line learning and an application to boosting", in *Conference on Learning Theory*, 1995, pp. 23-37.
- [7] L. Breiman, J. Friedman, C. J. Stone, e R. A. Olshen, *Classification and Regression Trees*. CRC press, 1984.
- [8] D. E. Knuth, J. H. Morris, Jr, e V. R. Pratt, "Fast pattern matching in strings", *SIAM Journal on Computing*, vol. 6, no. 2, pp. 323-350, 1977.
- [9] H. Wu, Q. Chen, e M. Yachida, "Face detection from color images using a fuzzy pattern matching method", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 6, pp. 557-563, 1999.
- [10] D. Singh, J. Pandey, e D. Chauhan, "Topology identification, bad data processing, and state estimation using fuzzy pattern matching", *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1570-1579, 2005.
- [11] J. Kim e M. Fan, "Detecção de pontos críticos no layout Pós-OPC usando ferramenta de verificação baseada em simulação completa de chip: Um estudo de caso com simulação de imagem aérea", em *Proceedings of SPTe*, vol. 5256, 2003.



292

Uso autorizado e licenciado limitado a: UNIVERSIDADE FEDERAL DO PARANAENSE. Download em 22.2023 de abril às 03:03:24 UTC do IEEE Xplore. As restrições são aplicáveis.

$$\min_{\alpha} : \frac{1}{2} \alpha^T Q \alpha - e^T \alpha + \lambda \alpha^T \alpha$$

- [12] E. Roseboom, M. Rossman, E.-e. Chang, e P. Hurat, "Automated fuU chip hotspot detection and removal flow for interconnect layers of cell-based designs," in *Proceedings of SPIE*, vol. 6521, 2007.
- [13] D. G. Drmanac, E. Liu, e L.-e. Wang, "Predicting variability in nanoscale lithography processes," in *IEEE/ACM Design Automation Conference (DAC)*, 2009, pp. 545-550.
- [14] D. Ding, X. Wu, J. Ghosh, e D. Z. Pan, "Machine learning based lithographic hotspot detection with critical-feature extraction and classification," in *IEEE International Conference on IC Design and Technology (ICIDT)*, 2009, pp. 219-222.
- [15] D. Ding, I. A. Torres, e D. Z. Pan, "High performance lithography hotspot detection with successively refined pattern identifications and machine learning," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, pp. 1621-1634, 2011.
- [16] J. -Y Wu, E. G. Pikus, A. Torres, e M. Marek-Sadowska, "Rapid layout pattern classification," em *IEEE/ACM Asia and South Pacific Design Automation Conference (ASPDAC)*, 2011, pp. 781-786.
- [17] Y.-T. Yu, G.-H. Lin, I. H.-R. Jiang, e e. Chiang, "Machine-learning based hotspot detection using topological classification and critical feature extraction," in *IEEE/ACM Design Automation Conference (DAC)*, 2013, pp. 671-676.
- [18] Z. Xiao, Y. Du, H. Tian, M. D. Wong, H. Yi, H.-S. P. Wong, e H. Zhang, "Verificação de modelo de auto-montagem dirigida (DSA)," in *IEEE/ACM Design Automation Conference (DAC)*, 2014, pp. 1-6.
- [19] J.-R. Gao, B. Yu, e D. Z. Pan, "Accurate lithography hotspot detection based on PCA-SVM classifier with hierarchical data clustering," in *Proceedings of SPIE*, vol. 9053, 2014.
- [20] B. Yu, I.-R. Gao, D. Ding, X. Zeng, e D. Z. Pan, "Accurate lithography hotspot detection based on principal component analysis-support vector machine classifier with hierarchical data clustering," *Journal of Micro/Nanolithography, MEMS, and MOEMS (1M3)*, vol. 14, no. 1, p. 011003, 2015.
- [21] H. Nosato, H. Sakanashi, E. Takahashi, M. Murakawa, T. Matsunawa, S. Maeda, S. Tanaka, e S. Mimotogi, "Hotspot prevention and detection method using an image-recognition technique based on higher-order local autocorrelation," *Journal of Micro/Nanolithography, MEMS, and MOEMS (1M3)*, vol. 13, no. 1, p. 011007, 2014.
- [22] T. Matsunawa, I.-R. Gao, B. Yu, e D. Z. Pan, "A new lithography hotspot detection framework based on AdaBoost classifier and simplified feature extraction," em *Proceedings of SPIE*, 2015.
- [23] A. B. Kahng, e.-H. Park, e X. Xu, "Fast dual graph-based hotspot detection," em *Proceedings of SPIE*, vol. 6349, 2006.
- [24] H. Yao, S. Sinha, e. Chiang, X. Hong, e Y. Cai, "Efficient process hotspot detection using range pattern matching," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2006, pp. 625- 632.
- [25] I. Xu, S. Sinha, e e. e. Chiang, "Accurate detection for process hotspots with vias and incomplete specification," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2007, pp. 839- 846.
- [26] Y.-T. Yu, y.-e. Chan, S. Sinha, I. H.-R. Jiang, e e. Chiang, "Accurate process-hotspot detection using critical design rule extraction," in *IEEE/ACM Design Automation Conference (DAC)*, 2012, pp. 1167-1172.
- [27] N. Ma, "Automatic IC hotspot classification and detection using pattern based clustering," dissertação de doutorado, Engenharia e Engenharia mecânica, Universidade da Califórnia, Berkeley, 2008.
- [28] J. Guo, E. Yang, S. Sinha, C. Chiang, e X. Zeng, "Improved tangent space based distance metric for accurate lithographic hotspot classification," in *IEEE/ACM Design Automation Conference (DAC)*, 2012, pp. 1173-1178.
- [29] E. M. Arkin, L. Mastigari, D. P. Huttenlocher, K. Kedem, e J. S. B. Mitchell, "An efficiently computable metric for comparing polygonal shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 3, pp. 209-216, 1991.
- [30] L. J. Latecki e R. Lakamper, "Shape similarity measure based on correspondence of visual parts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 10, pp. 1-6, 2000.
- [31] S. -Y. Lin, J.-Y. Chen, J.-C. Li, w.-Y. Wen, e S.-e. Chang, "A novel fuzzy matching model for lithography hotspot detection," in *IEEE/ACM Design Automation Conference (DAC)*, 2013, pp. 681-686.
- [32] D. Ding, B. Yu, I. Ghosh, e D. Z. Pan, "EPIC: Efficient prediction of IC manufacturing hotspots with a unified metaclassification formulation," in *IEEE/ACM Asia and South Pacific Design Automation Conference (ASPDAC)*, 2012, pp. 263-270.
- [33] J. Mitra, P. Yu, e D. Z. Pan, "RADAR: RET-aware detailed routing using fast lithography simulations," in *IEEE/ACM Design Automation Conference (DAC)*, 2005, pp. 369-372.
- [34] J. Yang, N. Rodriguez, O. Omedes, E. Gennari, y.-e. Lai, e Y. Mankad, "DRCPPlus in a router: automatic elimination of lithography hotspots using 2D pattern detection and correction," em *Proceedings of SPIE*, vol. 7641, 2010, p. 76410Q.
- [35] I.-R. Gao, H. Jawandha, P. Atkar, A. Walimbe, B. Baidya, O. Rizzo, e D. Z. Pan, "Auto-alinhamento de roteamento em conformidade com a dupla patente, com fluxo de verificação física no projeto," em *Proceedings of SPIE*, vol. 8684, 2013, p. 868408.
- [36] D. Ding, I.-R. Gao, K. Yuan, e D. Z. Pan, "AENEID: um roteador genérico detalhado de litografia amigável baseado no aprendizado de dados pós-RET e detecção de hotspots," in *IEEE/ACM Design Automation Conference (DAC)*, 2011, pp. 795-800.
- [37] S. I. Ward, D. A. Papa, Z. Li, e N. Sze, e I. Alpert, e E. Swartzlander, "Quantifying academic placer performance on custom designs," in *ACM International Symposium on Physical Design (ISPD)*, 2011, pp. 91-98.
- [38] S. Ward, D. Ding, e D. Z. Pan, "PADE: um placer de alto desempenho com extração e avaliação automática de dados através do aprendizado de dados de alta dimensão," in *IEEE/ACM Design Automation Conference (DAC)*, 2012, pp. 756-761.
- [39] S. Chou, M.-K. Hsu, e Y.-w. Chang, "Structure-aware placement for datapath-intensive circuit designs," in *IEEE/ACM Design Automation Conference (DAC)*, 2012, pp. 762-767.
- [40] H. Xiang, M. Cho, H. Ren, M. Ziegler, e R. Puri, "Network flow based datapath bit slicing," in *ACM International Symposium on Physical Design (ISPD)*, 2013, pp. 139-146.
- [41] N. Viswanathan, M. Pan, e e. Chu, "FastPlace 3.0: Um algoritmo rápido de posicionamento quadrático multinível com controle de congestionamento de posicionamento," em *IEEE/ACM Asia and South Pacific Design Automation Conference (ASPDAC)*, 2007, pp. 135-140.
- [42] D. Papa, e. Alpert, e. Sze, Z. Li, N. Viswanathan, G.-J. Nam, e I. L. Markov, "Physical synthesis with clock-network optimization for large systems on chips," *IEEE Micro*, vol. 31, no. 4, pp. 51-62, 2011.
- [43] M. Cho, H. Xiang, H. Ren, M. M. Ziegler, e R. Puri, "Latchplan ner: algoritmo de posicionamento da trava para projetos VLSI de alto desempenho orientados a dados," in *IEEE/ACM International Conference on Computer Aided Design (ICCAD)*, 2013, pp. 342-348.
- [44] S. Held e U. Schorr, "Post-routing latch optimization for timing closure," in *IEEE/ACM Design Automation Conference (DAC)*, 2014, pp. 1-6.
- [45] S. I. Ward, N. Viswanathan, N. Y. Zhou, e. e. Sze, Z. Li, e J. Alpert, e D. Z. Pan, "Clock power minimization using structure latch templates and decision tree induction," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2013, pp. 599-606.
- [46] "Calibre de correspondência de padrões," <http://www.mentor.com/products>.
- [47] "Synopsys IC Validator," <http://www.synopsys.com>.
- [48] "Cadence Virtuoso DFM," <http://www.cadence.com>.
- [49] "Anchor Semiconductor NanoScope," <http://www.anchorsemi.com>.
- [50] P. C. Mahalanobis, "On the generalized distance in statistics," *Proceedings do Instituto Nacional de Ciências (Calcutá)*, vol. 2, pp. 49-55, 1936.
- [51] Y. S. Abu-Mostafa, M. Magdon-Ismael, e H.-T. Lin, *Aprendendo com os Dados*. AMLBook, 2012.
- [52] P.-N. Tan, M. Steinbach, e Y. Kumar, *Introdução à mineração de dados*. Addison-Wesley, 2006.
- [53] P. J. Bickel, B. Li, A. B. Tsybakov, S. A. van de Geer, B. Yu, T. Valdes, e. Rivero, I. Fan, e A. van der Vaart, "Regularização em estatísticas," *Teste*, vol. 15, não. 2, pp. 271-344, 2006.
- [54] S. Arlot e A. Celisse, "A survey of cross-validation procedures for model selection," *Statistics Surveys*, vol. 4, pp. 40-79, 2010.

restrições são aplicáveis.