



República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(21) BR 102018009859-4 A2



(22) Data do Depósito: 15/05/2018

(43) Data da Publicação Nacional: 12/03/2019

(54) **Título:** MÉTODO E SISTEMA PARA OTIMIZAÇÃO COM BASE EM DADOS DE INDICADORES DE DESEMPENHO EM INDÚSTRIAS DE FABRICAÇÃO E PROCESSO

(51) **Int. Cl.:** G06Q 10/06.

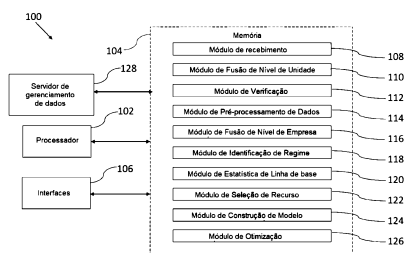
(52) **CPC:** G06Q 10/06.

(30) **Prioridade Unionista:** 15/05/2017 IN 201721009012.

(71) **Depositante(es):** TATA CONSULTANCY SERVICES LIMITED.

(72) **Inventor(es):** VENKATARAMANA RUNKANA; ROHAN PANDYA; RAJAN KUMAR; ANIRUDDHA PANDA; MAHESH MYNAM; SRI HARSHA NISTALA; PRADEEP RATHORE; JAYASREE BISWAS.

(57) **Resumo:** Trata-se de um sistema e método para realizar a otimização com base em dados de indicadores de desempenho de usinas de fabricação e processo. O sistema consiste em módulos para coletar e mesclar dados a partir de unidades de processamento industriais, pré-processando os dados para remover valores atípicos e omissão. Adicionalmente, o sistema gera saídas personalizadas a partir de dados e identifica variáveis importantes que afetam um determinado indicador de desempenho de processo. O sistema também constrói modelos preditivos para indicadores chave de desempenho que compreendem os recursos importantes e determina pontos de operação para otimizar os indicadores chave de desempenho com intervenção de usuário mínima. Em particular, o sistema recebe entradas a partir de usuários nos indicadores chave de desempenho a serem otimizados e notifica os usuários de saídas a partir de diversas etapas na análise que ajudam os usuários a gerenciar de modo eficaz a análise e tomar decisões operacionais adequadas.



Relatório Descritivo da Patente de Invenção para: "MÉTODO E SISTEMA PARA OTIMIZAÇÃO COM BASE EM DADOS DE INDICADORES DE DESEMPENHO EM INDÚSTRIAS DE FABRICAÇÃO E PROCESSO"

Descrição

Reivindicação de Prioridade

[001]Este pedido de patente reivindica a prioridade sobre o Pedido da Índia N°. 201721009012, depositada em 15 de maio de 2017. Todo o teor do pedido anteriormente mencionado está incorporado ao presente documento a título de referência.

Campo da Técnica

[002]As modalidades do presente documento referem-se, de modo geral, ao campo de análise de dados e, especificamente, a um sistema e um método para otimizar indicadores chave de desempenho de indústrias de fabricação e processo.

Antecedentes

[003]Indicadores, tais como produtividade, qualidade do produto, consumo de energia, tempo de atividade em porcentagem, níveis de emissão, etc., são usados para monitorar o desempenho das indústrias de fabricação e de usina de processos. Atualmente, as indústrias encaram o desafio de atingir metas de produção ambiciosas, o que minimiza o seu consumo de energia, atinge padrões de emissão e personaliza os seus produtos, enquanto lida com amplas

variações da qualidade de matéria-prima e outros parâmetros de influência, tais como temperatura ambiente, umidade, etc. As indústrias se esforçam para continuamente melhorar os seus indicadores de desempenho através da modulação de alguns parâmetros que são conhecidos como influência ou que afetam os mesmos. Isso é fácil quando um processo envolve um número limitado de variáveis. Entretanto, o maior dos processos industriais consiste em muitas unidades em série e/ou em paralelo, ou envolvem milhares de variáveis ou parâmetros. A identificação das variáveis que influenciam indicadores chave de desempenho (KPIs) e (seus) níveis de otimização em tais situações não é simples, e fazer o mesmo exige muito tempo e conhecimento específico. Os métodos de análise de dados, tais como técnicas de estatística, aprendizado por máquina e mineração de dados, têm potencial para solucionar esses problemas de otimização complexos e podem ser usados para analisar dados industriais e constatarem novos regimes de operação.

[004]A identificação de variáveis relevantes que afetam os KPIs é um desafio associado à análise de dados de processo. Isso ocorre devido ao grande número de variáveis em processos industriais e interações complexas não lineares entre os mesmos. Há várias técnicas de seleção variáveis (ou recursos), mas nenhuma técnica de seleção variável única é capaz de identificar todas as variáveis relevantes,

particularmente, em processos industriais complexos. Portanto, há uma necessidade de uma técnica de seleção variável melhor que seja capaz de selecionar as variáveis mais importantes.

[005]Além disso, em todos os métodos que descrevem a aplicação de análise de dados para industriais de processo de fabricação, o foco é limitado à visualização dos KPIs, outras variáveis de interesse e resultados dos modelos preditivos, e/ou que fornecem recomendações de processo ao usuário final. Várias outras saídas, tais como faixas de variáveis que correspondem às faixas desejadas e não desejadas de KPIs, faixas de KPIs em níveis de vazão diferentes, etc. são de grande ajuda para os usuários finais na tomada de decisão e não aparecem em qualquer um dentre os métodos existentes.

Sumário

[006]A seguir, é apresentado um sumário simplificado de algumas modalidades da divulgação a fim de fornecer uma compreensão simples das modalidades. Este sumário não é uma extensão da visão geral das modalidades. Não se pretende identificar elementos-chave/críticos das modalidades ou delinear o escopo das modalidades. O único propósito é apresentar algumas modalidades em uma forma simplificada como um prelúdio à descrição mais detalhada que é apresentada abaixo.

[007]Em vista do supracitado, uma modalidade do presente documento fornece um sistema e método para analisar uma pluralidade de dados de uma ou mais unidades de processamento industriais para otimizar os indicadores chave de desempenho da indústria.

[008]Em um aspecto, o que segue apresenta um sistema e um método para analisar uma pluralidade de dados de uma ou mais unidades de processamento industriais para otimizar os indicadores chave de desempenho da indústria. O sistema compreende uma memória com instruções, pelo menos um processador acoplado de maneira comunicativa com a memória, uma pluralidade de interfaces e uma pluralidade de módulos. Um módulo de recebimento é configurado para receber a pluralidade de dados de uma ou mais unidades de processamento industriais, em que a pluralidade de dados compreende características de matérias-primas, características de produtos intermediários, subprodutos e produtos finais, parâmetros de processo e condição de equipamentos do processo. Um módulo de fusão de nível de unidade é configurado para mesclar a pluralidade de dados recebidos para obter um conjunto de dados por unidade de cada uma dentre uma ou mais unidades de processamento industriais, em que o conjunto de dados por unidade de cada unidade de processamento compreende uma frequência de amostragem desejada. Um módulo de verificação é configurado para

verificar conjunto de dados por unidade mesclado de um ou mais unidades de processamento industriais, em que a presença de valores indesejados, porcentagem de disponibilidade, desvio padrão e faixa interquartil de todas as variáveis da unidade de processamento são calculados. Um módulo de pré-processamento de dados é configurado para pré-processar a pluralidade de dados verificada para obter um conjunto de dados pré-processado de cada uma dentre uma ou mais unidades de processamento industriais, em que o pré-processamento é um processo interativo que compreende as etapas de remoção de valor atípico, imputação de valores faltantes e agrupamento. Um módulo de fusão de nível de empresa é configurado para integrar os dados pré-processados de cada uma dentre uma ou mais unidades de processamento industriais com um ou mais valores de variáveis simuladas de um ou mais modelos com base em física e uma ou mais entradas de domínio do usuário para obter um conjunto de dados de nível de empresa, em que o conjunto de dados por unidades é mesclado e sincronizado levando-se em conta os intervalos de tempo devido a tempos de permanência em diversas unidades, tempos de transporte entre uma ou mais unidades de processamento industriais e tempo de resposta de um ou mais sensores das unidades de processamento. Um módulo de identificação de regime é configurado para identificar um ou mais regimes de operação com o uso de uma ou mais técnicas de agrupamento no

conjunto de dados de nível de empresa, em que uma ou mais técnicas de agrupamento compreendem agrupamento com base em distância, agrupamento com base em densidade e agrupamento hierárquico. Um módulo de estatística de linha de base é configurado para determinar faixas de uma ou mais variáveis que correspondem aos KPIs do conjunto de dados de nível de empresa. A determinação de faixa tem por base estatísticas de linha de base predefinidas e um ou mais regimes de operação, em que as faixas determinadas de uma ou mais variáveis estão sendo usadas para gerar uma ou mais plotagens de KPIs durante o período de tempo em que a análise está sendo executada. Um módulo de seleção de recurso é configurado para selecionar um ou mais recursos do conjunto de dados de nível de empresa para obter o superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa, em que a seleção de recurso é realizada em todos os conjuntos de dados por regime, bem como no conjunto de dados de nível de empresa. Um módulo de construção de modelo é configurado para desenvolver um ou mais modelos preditivos para cada KPI, em que o um ou mais modelos preditivos é desenvolvido com o uso do conjunto de dados de nível de empresa e do superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa. Um módulo de otimização é configurado para otimizar pelo menos um KPI com base em um ou mais modelos preditivos e

constringir um ou mais KPIs com o uso de uma ou mais técnicas de otimização, em que uma ou mais técnicas de otimização compreendem pesquisa de gradiente, programação linear, programação de objetivo, anelamento simulado e algoritmos evolucionários.

[009]Em outro aspecto, a seguir é apresentado um método para analisar uma pluralidade de dados de uma ou mais unidades de processamento industriais para otimizar os indicadores chave de desempenho da indústria. O método compreende as etapas de receber a pluralidade de dados de uma ou mais unidades de processamento industriais, em que a pluralidade de dados compreende características de matérias-primas, características de produtos intermediários, subprodutos e produtos finais, parâmetros de processo e condição de equipamentos do processo, que mesclam a pluralidade de dados recebidos para obter um conjunto de dados por unidade de cada uma dentre uma ou mais unidades de processamento industriais, verificar o conjunto de dados por unidade mesclado de um ou mais unidades de processamento industriais, em que a presença de valores absurdos, porcentagem de disponibilidade, desvio padrão e faixa interquartil de todas as variáveis da unidade de processamento são calculados, pré-processar a pluralidade de dados verificada para obter um conjunto de dados pré-processado de cada uma dentre uma ou mais unidades de

processamento industriais, em que o pré-processamento é um processo interativo que compreende as etapas de remoção de valor atípico, imputação de valores faltantes e agrupamento, integração do conjunto de dados pré-processados de cada uma dentre uma ou mais unidades de processamento industriais com um ou mais valores de um ou mais modelos com base em física, e uma ou mais entradas de domínio a partir do usuário para obter um conjunto de dados de nível de empresa, em que o conjunto de dados por unidades são mesclados e sincronizados levando-se em conta os intervalos de tempo devido a tempos de permanência em diversas unidades, tempos de transporte de materiais de uma ou mais unidades de processamento industriais e tempo de resposta de um ou mais sensores das unidades de processamento, identificar um ou mais regimes de operação com o uso de uma ou mais técnica de agrupamento no conjunto de dados de nível de empresa, em que uma ou mais técnicas de agrupamento compreendem agrupamento com base em distância, agrupamento com base em densidade e agrupamento hierárquico, determinar faixas de uma ou mais variáveis que correspondem aos KPIs do conjunto de dados de nível de empresa com base em estatísticas de linha de base predefinidas e o um ou mais regimes de operação, em que as faixas determinadas de uma ou mais variáveis está sendo usada para gerar uma ou mais plotagens de KPIs durante o período de tempo de análise que está sendo realizado, selecionar um

ou mais recursos do conjunto de dados de nível de empresa para obter um superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa, em que a seleção de recurso é realizada em todos os conjuntos de dados por regime, bem como no conjunto de dados de nível de empresa, desenvolver um ou mais modelos preditivos para cada KPI, em que o um ou mais modelos preditivos usam conjunto de dados de nível de empresa e o superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa e otimizar pelo menos um KPI com base em um ou mais modelos preditivos e restringir em um ou mais KPIs como uso de uma ou mais técnicas de otimização, em que uma ou mais técnicas de otimização compreendem pesquisa de gradiente, programação linear, programação de objetivo, anelamento simulado e algoritmos evolucionários.

[0010] Em ainda um outro aspecto, a modalidade do presente documento fornece um ou mais meio de armazenamento de informações legível por máquina não-transitório que compreende uma ou mais instruções, que quando executadas por um ou mais processadores de hardware realizam ações que compreendem receber uma pluralidade de dados de uma ou mais unidades de processamento industriais, em que a pluralidade de dados compreende características de matérias-primas, características de produtos intermediários, subprodutos e produtos finais, parâmetros de processo e condição de

equipamentos do processo, que mesclam a pluralidade de dados recebidos para obter um conjunto de dados por unidade de cada uma dentre uma ou mais unidades de processamento industriais, verificar o conjunto de dados por unidade mesclado de um ou mais unidades de processamento industriais, em que a presença de valores absurdos, porcentagem de disponibilidade, desvio padrão e faixa interquartil de todas as variáveis da unidade de processamento são calculados, pré-processar a pluralidade de dados verificada para obter um conjunto de dados pré-processado de cada uma dentre uma ou mais unidades de processamento industriais, em que o pré-processamento é um processo interativo que compreende as etapas de remoção de valor atípico, imputação de valores faltantes e agrupamento, integração do conjunto de dados pré-processados de cada uma dentre uma ou mais unidades de processamento industriais com um ou mais valores de um ou mais modelos com base em física, e uma ou mais entradas de domínio a partir do usuário para obter um conjunto de dados de nível de empresa, em que o conjunto de dados por unidades são mesclados e sincronizados levando-se em conta os intervalos de tempo devido a tempos de permanência em diversas unidades, tempos de transporte de materiais de uma ou mais unidades de processamento industriais e tempo de resposta de um ou mais sensores das unidades de processamento, identificar um ou mais regimes de operação

com o uso de uma ou mais técnica de agrupamento no conjunto de dados de nível de empresa, em que uma ou mais técnicas de agrupamento compreendem agrupamento com base em distância, agrupamento com base em densidade e agrupamento hierárquico, determinar faixas de uma ou mais variáveis que correspondem aos KPIs do conjunto de dados de nível de empresa com base em estatísticas de linha de base predefinidas e o um ou mais regimes de operação, em que as faixas determinadas de uma ou mais variáveis está sendo usada para gerar uma ou mais plotagens de KPIs durante o período de tempo de análise que está sendo realizado, selecionar um ou mais recursos do conjunto de dados de nível de empresa para obter um superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa, em que a seleção de recurso é realizada em todos os conjuntos de dados por regime, bem como no conjunto de dados de nível de empresa, desenvolver um ou mais modelos preditivos para cada KPI, em que o um ou mais modelos preditivos usam conjunto de dados de nível de empresa e o superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa e otimizar pelo menos um KPI com base em um ou mais modelos preditivos e restringir em um ou mais KPIs como uso de uma ou mais técnicas de otimização, em que uma ou mais técnicas de otimização compreendem pesquisa de gradiente, programação linear, programação de objetivo, anelamento simulado e algoritmos

evolucionários.

[0011] Deve ser observado pelos especialistas no assunto que qualquer diagrama de blocos do presente documento representa vistas conceituais dos sistemas ilustrativos que incorporam os princípios da matéria presente. Similarmente, será observado que quaisquer fluxogramas, diagramas de fluxo, diagramas de transição de estado, pseudocódigo e similares representam vários processos que podem ser substancialmente representados em meio legível por computadores e, então, executados por um dispositivo de computação ou processador, tanto se o dispositivo de computação ou processador esteja ou não explicitamente mostrado.

Breve Descrição dos Desenhos

[0012] As modalidades do presente documento serão melhor compreendidas a partir da seguinte descrição detalhada em referência aos desenhos, em que:

[0013] A Figura 1 ilustra um sistema para analisar uma pluralidade de dados de uma ou mais unidades de processamento industriais para otimizar os indicadores chave de desempenho da indústria de acordo com uma modalidade da presente divulgação;

[0014] A Figura 2 é um esquema de um usina de processo ou manufatura de acordo com uma modalidade da presente divulgação;

[0015] A Figura 3 é um esquema que mostra as etapas no método para otimizar KPIs de acordo com uma modalidade da presente divulgação;

[0016] As Figuras 4(a) e 4(b) são um fluxograma que representa pré-processamento de dados com o uso de valor atípico e imputação de técnicas de acordo com uma modalidade da presente divulgação;

[0017] A Figura 5 é um esquema das entradas e saídas da etapa de pré-processamento de dados de acordo com uma modalidade da presente divulgação;

[0018] A Figura 6 é um esquema das entradas e saídas da integração de nível de empresa de acordo com uma modalidade da presente divulgação;

[0019] A Figura 7 é um esquema das entradas e saídas das estatísticas de linha de base e do regime de identificação de acordo com uma modalidade da presente divulgação;

[0020] A Figura 8 é um fluxograma da seleção de recurso de acordo com uma modalidade da presente divulgação;

[0021] As Figuras 9(a) e 9(b) são um fluxograma da modelo de construção e discriminação de acordo com uma modalidade da presente divulgação;

[0022] A Figura 10 é um esquema das entradas e saídas do modelo de construção e discriminação de acordo com uma modalidade da presente divulgação;

[0023] A Figura 11 é um esquema das entradas e saídas da otimização de acordo com uma modalidade da presente divulgação; e

[0024] As Figuras 12(a) e 12(b) ilustram um método para analisar uma pluralidade de dados de uma ou mais unidades de processamento industriais para otimizar os indicadores chave de desempenho da indústria de acordo com uma modalidade da presente divulgação.

Descrição Detalhada das Modalidades

[0025] As modalidades do presente documento e os diversos recursos e detalhes vantajosos do mesmo são explicados mais completamente em referência às modalidades não limitadoras que são ilustradas nos desenhos em anexo e detalhadas na descrição a seguir. Os exemplos usados no presente documento são destinados a meramente facilitar a compreensão das formas em que as modalidades do presente documento podem ser praticadas e, além disso, possibilitar que os especialistas no assunto pratiquem as modalidades do presente documento. Consequentemente, os exemplos não devem ser interpretados como limitadores do escopo das modalidades do presente documento.

[0026] Em referência à Figura 1, um sistema 100 para analisar uma pluralidade de dados de uma ou mais unidades de processamento industriais para otimizar os indicadores chave de desempenho da indústria é apresentado. O sistema 100

compreende um processador 102, uma memória 104 acoplada de modo comunicativo ao processador 102, uma pluralidade de interfaces 106, um módulo de recebimento 108, um módulo de fusão de nível de unidade 110, um módulo de verificação 112, um módulo de pré-processamento de dados 114, um módulo de fusão de nível de empresa 116, um módulo de identificação de regime 118, um módulo de estatística de linha de base 120, um módulo de seleção de recurso 122, um módulo de construção de modelo 124, um módulo de otimização 126 e um servidor de gerenciamento de dados 128.

[0027] Na modalidade preferencial, a memória 104 contém instruções que são legíveis pelo processador 102. A pluralidade de interfaces 106 compreende interface de usuário gráfica, interface de servidor, uma interface de modelo com base em física e uma interface de resolução. A interface de usuário gráfica é usada para receber entradas, tais como os KPIs de interesse e o período de tempo de análise do usuário, e encaminhar os mesmos à pluralidade de módulos. A interface de servidor encaminha a solicitação de dados recebida de um dentre uma pluralidade de módulos ao servidor de gerenciamento de dados 128 e aos dados recebidos do servidor de gerenciamento de dados 128 à pluralidade de módulos. A interface de modelo com base em física envia o conjunto de dados recebidos integrados de um dentre uma pluralidade de módulos após a fusão de nível de empresa para

os modelos com base em física, se houver, disponível para o processo industrial, receber os valores de variáveis simuladas dos modelos com base em física e encaminhar os mesmos a um dentre uma pluralidade de módulos.

[0028] Na modalidade preferencial, um módulo de recebimento 108 é configurado para receber a pluralidade de dados de uma ou mais unidades de processamento industriais, em que a pluralidade de dados compreende características de matérias-primas, características de produtos intermediários, subprodutos e produtos finais, parâmetros de processo e condição de equipamentos do processo.

[0029] Em referência às Figuras 2 e 3, como exemplos, um esquema de uma empresa industrial hipotética em que a maioria das empresas de processo e manufatura consiste em várias unidades em série ou em paralelo. A empresa consiste em 8 unidades de processo que produzem dois produtos, isto é, A e B, para produzir o produto A, o fluxo de materiais ocorre através da seguinte sequência de operações: (Unidade nº1, Unidade nº2, Unidade nº3) → Unidade nº4 → Unidade nº5 → Unidade nº6. Similarmente, para produzir o produto B, o fluxo de materiais ocorre através da seguinte sequência de operações: (Unidade nº1, Unidade nº2, Unidade nº3) → Unidade nº4 → Unidade nº7 → Unidade nº8. A fim de otimizar os KPIs relacionados à produção do produto A, a dita qualidade do produto A ou energia consumida por massa de unidade do

produto A produzido, dados de todas as unidades envolvidas na sequência de operacional devem ser considerados. Similarmente, a fim de otimizar os KPIs relacionados à produção do produto B, os dados de todas as unidades envolvidas na sequência de operacional devem ser considerados. Análise dos dados de nível de empresa em vez de dados de nível de unidade podem render perspectivas melhores nas operações da empresa. A Figura 3 mostra que para cada unidade de processo de N unidades de processo, os dados são coletados a partir de várias fontes, tal como Planejamento de Recursos Empresariais (ERP), Sistema de Controle Distribuído (DCS) e Sistema de Gerenciamento de Informações Laboratoriais (LIMS).

[0030] Na modalidade preferencial, o módulo de fusão de nível de unidade é configurado para mesclar a pluralidade de dados recebidos para obter um conjunto de dados por unidade de cada uma dentre uma ou mais unidades de processamento industriais, em que o conjunto de dados por unidade de cada unidade de processamento compreende uma frequência de amostragem desejada. No processo de mescla, a uma ou mais variáveis de todos os arquivos ou conjunto de dados é mesclada como por observação específica de ID que corresponde à frequência de amostragem, por exemplo, dados no caso de dados diários, horas no caso de dados de hora, etc. Se a frequência de amostragem é inconsistente em relação

à vários arquivos/conjunto de dados, valores de variáveis são calculados em média sempre que possível. Se o cálculo de média não for possível, os mesmos dados são usados em relação a, por exemplo, quando análises de horas está para ser realizada e apenas quando dados diários estão disponíveis, o valor de dados diários é usado para todas as horas no dia específico. No final do processo, o conjunto de dados por unidades com linhas que corresponde ao ID de observação e as colunas que correspondem a todas as variáveis na unidade de processo são obtidos.

[0031] Na modalidade preferencial, o módulo de verificação é configurado para verificar conjunto de dados por unidade mesclado de um ou mais unidades de processamento industriais, em que a presença de valores indesejados, porcentagem de disponibilidade, desvio padrão e faixa interquartil de todas as variáveis da unidade de processamento são calculados. A verificação da qualidade de dados é realizada no conjunto de dados por unidades obtido para cada uma dentre as unidades de processo. Os mapas de omissão que mostram a porcentagem e o padrão da disponibilidade das variáveis também são criados para cada unidade de processo. As métricas de qualidade dos dados e os mapas de omissão são enviados como saídas ao usuário por meio da interface de usuário. Dependendo da disponibilidade dos dados, o usuário pode decidir se procede ou não com o

restante da análise. O usuário também pode sugerir a deleção das mesmas variáveis com disponibilidade muito baixa antes de executar o restante das etapas.

[0032] Em referência às Figuras 4 (a), 4(b) e 5, em que o módulo de pré-processamento de dados 114 é configurado para pré-processar a pluralidade de dados verificada para obter um conjunto de dados pré-processado de cada uma dentre uma ou mais unidades de processamento industriais, em que o pré-processamento é um processo iterativo que compreende as etapas de remoção de valor atípico, imputação de valores faltantes e agrupamento. Variáveis com uma disponibilidade de porcentagem de pelo menos setenta por cento são consideradas para pré-processamento, embora essa condição esteja relaxada para variáveis de material, tal como matérias-primas, produto intermediário e características de produto final à medida que a omissão desses tipos de variáveis possa ocorrer devido ao número inferior de amostras visto que a análise laboratorial é genericamente executada apenas em intervalos periódicos.

[0033] As variáveis de material com menos disponibilidade do que a disponibilidade desejada e que não seguem nenhum padrão específico na omissão são descartadas do conjunto de dados. Uma análise de valor nominal atípico é inicialmente executada a fim de detectar e remover os valores atípicos o conjunto de dados, o que inclui valores

inconsistentes oriundos devido à falha/defeito de instrumento. No caso da produção de uma unidade ser zero, todas as variáveis para a unidade para esse período de tempo são negligenciadas. As variáveis são, então, categorizadas em vários subconjuntos com base na disponibilidade de porcentagem da variável. Enquanto imputação multivariada é usada para parâmetros de processo e variáveis de característica de material não sazonal, imputação em série de tempo é usada para variáveis sazonais de qualidade. Após a omissão em todas as variáveis em imputação apropriada, o agrupamento é realizado no conjunto de dados por unidade para identificar agrupamentos, se houver algum presente nos dados. Esses agrupamentos são representativos de regimes diferentes de operação. Cada conjunto de dados por unidade é, então, dividido em conjuntos de dados diferentes com base nos agrupamentos identificados. Os conjuntos de dados divididos são obtidos através das etapas de remoção de valor atípico e imputação conforme mostrado nas Figuras 4(a) e 4(b).

[0034] Na modalidade preferencial, o processo iterativo da remoção de valor atípico, imputação e agrupamento é interrompido quando o número de agrupamentos e o número de pontos de dados em cada agrupamento não é alterado. Os conjuntos de dados pré-processados por unidade são obtidos no final dessa etapa. Para cada variável, o

número/porcentagem dos valores atípicos removidos, a técnica usada para imputação, e média, desvio mediano e padrão antes e depois do pré-processamento são apresentadas ao usuário como saídas. A lista de variáveis descartadas também é apresentada ao usuário. O usuário também é dotado de opção de visualização das tendências de variáveis pré-processadas e originais.

[0035] Na modalidade preferencial, em referência à Figura 6, o módulo de fusão de nível de empresa 116 é configurado para integrar os dados pré-processados de cada uma dentre uma ou mais unidades de processamento industriais com um ou mais valores de variáveis simuladas de um ou mais modelos com base em física e uma ou mais entradas de domínio do usuário para obter um conjunto de dados de nível de empresa, em que o conjunto de dados por unidades é mesclado e sincronizado levando-se em conta os intervalos de tempo devido a tempos de permanência em diversas unidades, tempos de transporte entre uma ou mais unidades de processamento industriais e tempo de resposta de um ou mais sensores das unidades de processamento. Se o tempo de transporte entre as duas unidades de processo é maior do que a frequência de amostragem dos dados, então, os IDs de observação para uma dentre as unidades de processo é alterado através de número apropriado de unidades de tempo antes da integração. Por exemplo, se a frequência de amostragem é diária e leva 2

dias para o material percorrer da unidade de processo A à unidade de processo B, então, todos os IDs de observação no conjunto de dados do processo A são alterados por 2 dias antes de mesclar os conjuntos de dados de ambos os processos.

[0036] Na modalidade preferencial, qualquer unidade de processo específica pode ser considerada como a linha de base para mescla dos dados de todas as unidades de processo. Tipicamente, a unidade de processo na qual os KPIs de interesse são calculados é considerada como sendo a unidade de linha de base para integração dos dados. Nesse caso, o mesmo produto intermediário está saindo de duas ou mais unidades de processo diferentes, então, as variáveis de operação de todas as tais unidades de processo são consideradas para análise. Entretanto, em vez de usar características de material (análise de tamanho, análise química, etc.) de todas as unidades de processo em que o produto intermediário é gerado, as características em média ponderada são usadas. Os pesos poderiam ser quantidades do produto intermediário gerado de cada uma dentre as unidades de processo ou as quantidades do produto intermediário consumido na unidade de processo subsequente.

[0037] Uma vez que o conjunto de dados de nível de empresa é preparado, o mesmo é encaminhado aos modelos com base em física, se qualquer um estiver disponível para o processo industrial por meio da interface de modelo com base

em física para cálculo de variáveis simuladas. Esses são parâmetros que podem ter um impacto nos KPIs, mas não podem ser diretamente medidos no processo. Exemplos de variáveis simuladas são temperatura em zona temperatura alta (>1500 °C) de um forno, concentração do produto intermediário em um reator, etc. Os parâmetros simulados são enviados de volta ao módulo de fusão de nível de empresa e são adicionados ao conjunto de dados de nível de empresa para obter o conjunto de dados integrado para análise adicional. Saídas da integração de nível de empresa compreendem faixa, média, desvio padrão e mediano de todas as variáveis, e a lista de parâmetros estimados e simulados.

[0038] Na modalidade preferencial, o módulo de identificação de regime 118 é configurado para identificar um ou mais regimes de operação com o uso de uma ou mais técnicas de agrupamento no conjunto de dados de nível de empresa, em que uma ou mais técnicas de agrupamento compreendem agrupamento com base em distância, agrupamento com base em densidade e agrupamento hierárquico.

[0039] Na modalidade preferencial, o módulo de estatística de linha de base 120 é configurado para determinar faixas de uma ou mais variáveis dos KPIs do conjunto de dados de nível de empresa, com base nas estatísticas de linha de base predefinidas e o um ou mais regimes de operação, em que as faixas de uma ou mais

variáveis está sendo usada para gerar uma ou mais plotagens de KPIs durante o período de tempo de análise que está sendo executado. Estatística de linha de base, tais como porcentagem do tempo de KPIs estão nas faixas desejadas e não desejadas, as faixas de variáveis que correspondem às faixas desejadas e não desejadas de KPIs, as faixas de KPIs em níveis de vazão diferentes, e os coeficiente de correlação entre os KPIs e outras variáveis no conjunto de dados integrado são calculados e notificados ao usuário. É fornecido ao usuário a opção de gerar plotagens de tendência e plotagens de caixa de KPIs e todas as variáveis no conjunto de dados integrado durante o período de tempo no qual a análise está sendo executada. O usuário também pode gerar plotagens de dispersão entre os KPIs e as variáveis de interesse. Todas as variáveis no conjunto de dados integrado são binadas em vários intervalos entre seus valores mínimos e máximos. Os valores de KPI que corresponde a cada reservatório de cada variável são separados e sua média é calculada. Os valores de média de KPI que correspondem aos reservatórios/intervalo de todas as variáveis estão representados na forma de um mapa de calor e notificados ao usuário.

[0040] Na modalidade preferencial, o módulo de seleção de recurso 122 é configurado para selecionar um ou mais recursos do conjunto de dados de nível de empresa para

obter a superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa, em que a seleção de recurso é realizada em todos os conjuntos de dados por regime, bem como no conjunto de dados de nível de empresa. O conjunto de dados integrado é dividido em dois ou mais conjuntos de dados dependendo do número de regimes identificado durante a etapa de identificação de regime.

[0041] Seria contemplado que uma abordagem de seleção de recurso em dois estágios conforme mostrado na Figura 8 fosse usada para selecionar os recursos importantes. No primeiro estágio, recursos importantes são obtidos a partir de vários métodos de seleção de recurso. Esse estágio envolve o ajuste dos parâmetros disponíveis nos algoritmos de seleção de recurso e validação cruzada por k vezes para obter recursos importantes. Os métodos de seleção de recurso poderiam ser métodos com base em modelo, tais como floresta aleatória, ranhuras de regressão adaptativa multivariada, análise de componentes principais supervisionada, regressão gradual e regressão vetorial de suporte, métodos sem base em modelo, tais como mineração associada e agrupamento em série de tempo. No segundo estágio, as listas de recursos importantes obtidos de técnicas de seleção de recurso individuais são combinadas para obter um único 'superconjunto' de recursos importantes. Isso é obtido através da pontuação dos recursos principais identificados

25 por todas as técnicas com o uso de método de pontuação de média geométrica. A pontuação para o recurso '*i*' é calculada da seguinte forma:

$$\text{Pontuação}_i = \left(\prod_k R_{i,k} \right)^{1/n_i}$$

[0042] n_i é a frequência ou número de métodos que selecionaram o *i*-ésimo recurso; e

[0043] $R_{i,k}$ é a classificação do recurso *i* no *k*-ésimo método.

[0044] O superconjunto de recursos importantes junto com sua pontuação de importância em relação ao KPIs para os conjuntos de dados por regime e o conjunto de dados integrado são notificados ao usuário. É fornecido ao usuário a opção de adicionar outros recursos ou deletar recursos existentes dos superconjuntos. Para cada conjunto de dados, plotagens de coordenada paralela também são exibidas ao usuário.

[0045] Em referência às Figuras 9(A) e 9(b), o módulo de construção de modelo 124 do sistema 100 é configurado para desenvolver um ou mais modelos preditivos para cada KPI no conjunto de dados de treinamento, em que o um ou mais modelos preditivos com o uso do conjunto de dados de nível de empresa e do superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa. Seria

contemplado que uma abordagem de construção de modelo de três etapas seja usada. A primeira etapa envolve modelos preditivos de construção que usam algoritmos de construção de modelo básico. O um ou mais modelos preditivos compreende regressão gradual, regressão de componente principal, ranhuras de regressão adaptável multivariada, regressão de componente independente, regressão de laço, krigagem, floresta aleatória, quadrados mínimos parciais, árvores impulsionadas gradientes, modelo linear generalizado, máquinas de vetor de suporte linear e não lineares e redes neurais artificiais. A segunda etapa envolve ajustar os parâmetros do modelo de construção a fim de otimizar o desempenho de previsão dos modelos. O desempenho de previsão dos modelos é avaliado com o uso do conjunto de dados de teste e é expresso nos termos de erro médio quadrático (RMSE) de previsão, erro médio absoluto (MAE) de previsão, critério de informações akaike (AIC), critério de informações akaike de correção (AICc) e o critério de informações Bayesiana (BIC) e taxa de acertos (% de pontos com uma certa exatidão prevista) conforme mostrado na Figura 10. Seria contemplado que se em qualquer caso nenhum dos modelos preditivos alcançarem o RMSE e/ou MAE, é fornecido ao usuário a opção de voltar à seleção de recurso em que variáveis adicionais ou variáveis transformadas podem ser adicionadas ao superconjunto de variáveis importantes e repetirem a etapa

de construção de modelo.

[0046] A terceira etapa envolve discriminação de modelo e seleção em que para o conjunto de dados integrado e os conjuntos de dados por regime, os três principais modelos preditivos com valores de erro médio quadrático e erro médio absoluto são menores do que os valores especificados pelo usuário escolhidos. Uma pontuação de robustez (RS) é avaliada para os três principais modelos e usada para discriminação de modelo. Pelo menos dez mil pontos de dados que contêm valores de todas as variáveis compreendidas nos modelos são gerados aleatoriamente e usados para prever o KPI. A pontuação de robustez para cada modelo é, então, determinada com o uso,

$$RS = \frac{\text{Número de pontos de dados em que KPI está dentro de uma faixa desejada}}{\text{Número total de pontos de dados em que KPI é previsto}}$$

[0047] Os modelos preditivos em que a pontuação de robustez maior é maior do que 95% é selecionado para análise de sensibilidade e otimização. A análise de sensibilidade com base em variância é realizada para acessar a sensibilidade do KPI para mudanças de unidades nas variáveis no modelo. As pontuações de sensibilidade para cada uma das variáveis nos modelos são obtidas, com uma pontuação maior que indica uma mudança maior no valor do KPI com mudança de

unidade no valor da variável. Seria contemplado que se a pontuação de robustez para todos os três modelos preditivos for menor do que 95%, o usuário pode modificar o superconjunto de recursos importantes e repetir a etapa de construção de modelo.

[0048] Seria contemplado que o desempenho preditivo dos modelos é provável de diminuir com o tempo à medida que dados novos/futuros são obtidos para predição e uma opção de "autoaprendizagem" é fornecida ao usuário para melhorar a exatidão dos modelos preditivos. Para autoaprendizagem, dados originais usados para desenvolver os modelos e dados para o novo período de tempo são combinados e a etapa de construção de modelo é repetida no conjunto de dados combinados. A autoaprendizagem pode ser acionada automaticamente em uma base periódica (por exemplo, toda semana ou todo mês) ou pelo usuário com base nas medições estatísticas relacionadas aos modelos ou novo conjunto de dados. Medições estatísticas aos modelos poderiam ser métricas de desempenho, tal como erro médio quadrático, erro médio absoluto, critério de informações akaike, critério de informações akaike de correção, critério de informações de bayesiana ou taxa de acertos enquanto medições estatísticas relacionadas ao novo conjunto de dados poderiam ser desvio de porcentagem de novos dados dos dados originais ou distância multivariada entre conjunto de dados originais e

conjunto de dados novos.

[0049] Na modalidade preferencial, o módulo de otimização é configurado para otimizar pelo menos um KPI com base em um ou mais modelos preditivos e restringir um ou mais KPIs com o uso de uma ou mais técnicas de otimização, em que uma ou mais técnicas de otimização compreendem pesquisa de gradiente, programação linear, anelamento simulado e algoritmos evolucionários.

[0050] Em referência à Figura 11, um diagrama esquemático da otimização em que os KPIs a serem otimizados com restrições nas variáveis usadas nos modelos preditivos são considerados como entradas do usuário e os valores das variáveis que rendem níveis de otimização de KPIs são determinados. Quando qualquer um dos KPIs precisa ser otimizado, o problema é minimizar ou maximizar o KPI e a solução consiste em valores de variáveis que levam ao KPI mínimo/máximo. Quando dois ou mais KPIs precisam ser otimizados simultaneamente, o problema é minimizar a função de custo (por exemplo, função de custo = $0,6 \text{ KPI}_1 + 0,4 \text{ KPI}_2 - 1,2 \text{ KPI}_3$) e a solução consiste em um conjunto de pontos de operação pareto-ideais para o processo. A função de custo para otimização de KPI-múltiplo é construída com o uso dos pesos atribuídos a cada um dos KPIs pelo usuário. Várias técnicas de otimização, tal como pesquisa de gradiente, programação linear, programação de objetivo, anelamento

simulado e algoritmos evolucionários como algoritmos genéticos são usados. O problema de otimização é encaminhado aos resolvedores de otimização para algoritmos de otimização multiobjetivo ou de único objetivo, tal como com base em regra, com base em lógica difusa e resolvedores com base em gradiente através da interface de resolução. As soluções recebidas dos resolvedores são processadas e notificadas ao usuário. Saídas de usuário da etapa de otimização compreendem os valores de variáveis que rendem KPIs otimizados e os valores otimizados de KPIs, conjunto de pontos de operação pareto-ideais e os valores de KPIs nesses pontos, e a plotagem dos pontos de operação pareto-ideais.

[0051] Em referência à Figura 12(a) e 12(b), um método 400 para analisar uma pluralidade de dados de uma ou mais unidades de processamento industriais para otimizar os indicadores chave de desempenho (KPIs) da indústria.

[0052] Na etapa 402, em que o módulo de recebimento recebe a pluralidade de dados de uma ou mais unidades de processamento industriais, em que a pluralidade de dados compreende características de matérias-primas, características de produtos intermediários, subprodutos e produtos finais, parâmetros de processo, ambiente, demanda de mercado, disponibilidade de matérias-primas e condição de equipamentos do processo.

[0053] Na etapa 404, em que o módulo de fusão de

nível de unidade mescla a pluralidade de dados recebidos para obter um conjunto de dados por unidade de cada uma dentre uma ou mais unidades de processamento industriais, em que o conjunto de dados por unidade de cada unidade de processamento compreende uma frequência de amostragem desejada.

[0054] Na etapa 406, em que o módulo de verificação verifica o conjunto de dados por unidade mesclado de um ou mais unidades de processamento industriais, em que a presença de valores indesejados, porcentagem de disponibilidade, desvio padrão e faixa interquartil de todas as variáveis da unidade de processamento são calculados.

[0055] Na etapa 408, em que o módulo de pré-processamento de dados pré-processa a pluralidade de dados verificada para obter um conjunto de dados pré-processado de cada uma dentre uma ou mais unidades de processamento industriais, em que o pré-processamento é um processo interativo que compreende as etapas de remoção de valor atípico, imputação de valores faltantes e agrupamento. As saídas para o usuário do módulo de pré-processamento de dados compreendem lista de variáveis descartadas, número e porcentagem de valores atípicos removidos para cada variável, técnica usada para imputação de valores faltantes e cada variável, média, desvio padrão e mediano de cada variável antes e depois do pré-processamento, e plotagens de

tendência de todas as variáveis antes e depois do pré-processamento.

[0056] Na etapa 410, em que o módulo de fusão de nível de empresa integra os dados pré-processados de cada uma dentre uma ou mais unidades de processamento industriais com um ou mais valores de variáveis simuladas de um ou mais modelos com base em física e uma ou mais entradas de domínio do usuário para obter um conjunto de dados de nível de empresa, em que o conjunto de dados por unidades é mesclado e sincronizado levando-se em conta os intervalos de tempo devido a tempos de permanência em diversas unidades, tempos de transporte entre uma ou mais unidades de processamento industriais e tempo de resposta de um ou mais sensores das unidades de processamento. As saídas para o usuário do módulo de fusão de nível de empresa compreendem lista de parâmetros simulados, e a faixa, média, desvio padrão e mediano de todas as variáveis no conjunto de dados integrado.

[0057] Na etapa 412, o módulo de identificação de regime identifica um ou mais regimes de operação com o uso de uma ou mais técnicas de agrupamento no conjunto de dados de nível de empresa, em que uma ou mais técnicas de agrupamento compreendem agrupamento com base em distância, agrupamento com base em densidade e agrupamento hierárquico.

[0058] Na etapa 414, o módulo de estatística de linha de base determina faixas de uma ou mais variáveis

correspondentes aos KPIs do conjunto de dados de nível de empresa, com base nas estatísticas de linha de base predefinidas e o um ou mais regimes de operação, em que as faixas determinadas de uma ou mais variáveis está sendo usada para gerar uma ou mais plotagens de KPIs durante o período de tempo de análise que está sendo executado. As saídas para o usuário do módulo de estatística de linha de base compreende as porcentagens de KPIs de período de tempo que estão nas faixas desejadas e não desejadas, as faixas das variáveis que correspondem às faixas desejadas e não desejadas de KPIs, as faixas de KPIs em diferentes níveis de produtividade, coeficientes de correlação entre KPIs e outras variáveis, plotagens de tendência e plotagens de caixa de KPIs e outras variáveis, plotagens de dispersão entre KPIs e variáveis de interesse, e mapas de calor de valores de média de KPIs.

[0059] Na etapa 426, o módulo de seleção de recurso seleciona um ou mais recursos do conjunto de dados de nível de empresa para obter um superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa, em que a seleção de recurso é realizada em todos os conjuntos de dados por regime, bem como no conjunto de dados de nível de empresa. As saídas para o usuário do módulo de seleção de recurso compreendem o superconjunto de recursos e suas pontuações de importância para conjuntos de base de dados

integrados e por regime, e plotagens de coordenada paralelas dos recursos.

[0060] Na etapa 418, o módulo de modelo de construção desenvolve um ou mais modelos preditivos para cada KPI, em que o um ou mais modelos preditivos é usa conjunto de dados de nível de empresa e o superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa. As saídas para o usuário do módulo de discriminação e construção de modelo compreende métricas de desempenho para todos os modelos preditivos, três principais modelos preditivos desenvolvidos na base de RMSE e MAE, pontuação de robustez para os três principais modelos, pontuações de sensibilidade para todas as variáveis nos modelos robustos. Além disso, as saídas para o usuário do módulo de discriminação e construção de modelo também compreende plotagens de tendência dos valores atuais e previstos de KPI, plotagens de dispersão de valores atuais versus previstos de KPI e plotagens residuais de erro absoluto versus todas as variáveis nos modelos robustos.

[0061] Na etapa final 420, o módulo de otimização otimiza pelo menos um KPI com base em um ou mais modelos preditivos e restringir um ou mais KPIs com o uso de uma ou mais técnicas de otimização, em que uma ou mais técnicas de otimização compreendem pesquisa de gradiente, programação linear, anelamento simulado e algoritmos evolucionários. As

saídas para o usuário do módulo de otimização compreendem os valores de variáveis que rendem KPIs otimizados (pontos de operação pareto-ideais), valores otimizados de KPIs e plotagens dos pontos de operação pareto-ideais.

[0062] A descrição escrita descreve o assunto do presente documento para possibilitar que qualquer especialista na técnica produza e use as modalidades. O escopo das modalidades da matéria é definido pelas reivindicações e pode compreender outras modificações que ocorrerem aqueles que são especialistas na técnica. Tais outras modificações são destinadas a estarem dentro do escopo das reivindicações se tiverem elementos similares que não diferem da linguagem literal das reivindicações ou se compreenderem elementos equivalentes com diferenças insubstanciais da linguagem literal das reivindicações.

[0063] Um sistema e método para desempenhar otimização com base em dados dos indicadores de desempenho de usinas de fabricação e processo. O sistema consiste em módulos para coletar e mesclar dados das unidades de processamento industriais, pré-processar os dados para remover os valores atípicos e omissão. Além disso, o sistema gera saídas personalizadas de dados e identificar variáveis importantes que afetam um determinado indicador de processo de desempenho. O sistema também constrói modelos preditivos para indicadores chave de desempenho que compreendem os

recursos importantes e determina pontos de operação para otimizar os indicadores chave de desempenho com intervenção de usuário mínima. Em particular, o sistema recebe entradas dos usuários nos indicadores chave de desempenho para serem otimizados e notifica aos usuários das saídas de várias etapas na análise que auxilia os usuários a gerenciar efetivamente a análise e obter decisões operacionais apropriadas.

[0064] A modalidade da presente divulgação do presente documento é direcionado ao problema não solucionado de otimização de indicadores de desempenho para monitorar o desempenho de indústrias de fabricação e usina de processos, além de pré-processamento dos dados industriais recebidos da variedade de fontes que têm diferentes formatos e frequências de registro.

[0065] Entretanto, deve-se compreender que o escopo de proteção é estendido a tal programa e, além disso, a um meio legível por computador que tem uma mensagem no mesmo; tal meio de armazenamento legível por computador contém meios de código de programa para implementação de uma ou mais etapas do método, quando o programa é executado em um servidor ou dispositivo móvel ou dispositivo programável adequado. O dispositivo de hardware pode ser de qualquer tipo de dispositivo que pode ser programado, o que compreende, por exemplo, qualquer tipo de computador, como

um computador pessoal ou de serviço, ou similares, ou qualquer combinação dos mesmos. O dispositivo também pode compreender meios que poderia ser, por exemplo, meios de hardware, como por exemplo, um circuito integrado para aplicação específica (ASIC), uma matriz de portas programável em campo (FPGA), ou uma combinação de meios de hardware e software, por exemplo, um ASIC e um FPGA ou pelo menos um microprocessador e pelo menos uma memória com módulos de software localizados na mesma. Portanto, os meios podem compreender tantos meios de hardware quanto meios de software. As modalidades do método descritas no presente documento poderiam ser implementadas em hardware e software. O dispositivo também pode compreender meios de software. Alternativamente, as modalidades podem ser implementadas em dispositivos de hardware diferentes, por exemplo, com o uso de uma pluralidade de unidades de processamento central (CPUs).

[0066] A modalidades do presente documento pode compreender elementos de hardware e software. As modalidades que são implementadas em software compreendem, mas não se limitam a, firmware, software residente, microcódigo, etc. as funções realizadas pelos vários módulos descritos no presente documento podem ser implementadas em outros módulos ou combinações de outros módulos. Para os propósitos dessa descrição, uma mídia utilizável por computador ou legível

por computador pode ser qualquer aparelho que pode compreender, armazenar, comunicar, propagar ou transportar o programa para uso por ou em conexão com o sistema, o aparelho ou o dispositivo de execução de instruções.

[0067] A mídia pode ser um sistema (ou aparelho ou dispositivo) eletrônico, magnético, óptico, eletromagnético, infravermelho ou semicondutor ou uma mídia de propagação. Exemplos de uma mídia legível por computador compreendem uma memória semicondutora ou de estado sólido, fita magnética, um disquete de computador removível, uma memória de acesso aleatório (RAM), uma memória apenas leitura (ROM), um disco magnético rígido e um disco óptico. Os exemplos atuais de discos ópticos compreendem memória apenas leitura em disco compacto (CD-ROM), leitura/registro em disco compacto (CD-R/W) e disco de vídeo digital (DVD).

[0068] Um sistema de processamento de dados adequado para armazenar e/ou executar código de programa compreenderá pelo menos um processador (processador 28) acoplado direta ou indiretamente a elementos de memória através de um barramento de sistema. Os elementos de memória podem compreender memória local empregada durante execução real do código de programa, armazenamento bruto e memórias em cache que fornecem armazenamento temporário de pelo menos algum código de programa a fim de reduzir o número de vezes que o código precisa ser recuperado de um armazenamento bruto

durante a execução.

[0069] Dispositivos de entrada/saída (I/O) (que compreendem, mas não se limitam à teclados, telas, dispositivos de apontadores, etc.) podem ser acoplados ao sistema tanto diretamente quanto através de interversão dos controladores de I/O. Os adaptadores de rede também podem ser acoplados ao sistema para possibilitar que o sistema de processamento de dados se torne acoplado a outros sistemas de processadores de dados ou impressoras remotas ou dispositivos de armazenamento através de redes privadas ou públicas de intervenção. Modens, modem por cabo e cartões de Ethernet são apenas alguns dos tipos atualmente disponíveis de adaptadores de rede.

[0070] Um ambiente de hardware representativo para praticar as modalidades pode compreender uma configuração de hardware de um sistema de computação/manuseio de informações de acordo com as modalidades do presente documento. O sistema do presente documento compreende pelo menos um processador ou unidade de processamento central (CPU). Os CPUs são interconectados por meio do barramento de sistema para vários dispositivos, tal como uma memória de acesso aleatório (RAM), memória somente de leitura (ROM) e um adaptador de entrada/saída (I/O). O adaptador de I/O pode ser conectado aos dispositivos periféricos, tal como unidades de disco e drives de fita, ou outros dispositivos de armazenamento de

programa que são legíveis pelo sistema. O sistema pode ler as instruções da invenção nos dispositivos de armazenamento de programa e seguir essas instruções para executar a metodologia das modalidades do presente documento.

[0071] O sistema compreende adicionalmente um adaptador de interface de usuário que conecta um teclado, mouse, alto-falante, microfones, e/ou outros dispositivos de interface de usuário, tal como um dispositivo de tela sensível ao toque (não mostrado) para o barramento para reunir a entrada de usuário. Adicionalmente, um adaptador de conexão conecta o barramento a uma rede de processamento de dados, e um adaptador de tela conecta o barramento a um dispositivo de tela que pode ser incorporado como um dispositivo de saída, tal como um monitor, impressora ou transmissor, por exemplo.

[0072] A descrição anterior foi apresentada em referência às várias modalidades. Especialistas na técnica e na tecnologia na qual essa aplicação pertence apreciará que as alterações e mudanças nos métodos e nas estruturas descritos de operação podem ser praticados sem se afastar significativamente do princípio, essência e escopo.

REIVINDICAÇÕES

1. Método implantado por computador para analisar uma pluralidade de dados a partir de uma ou mais unidades de processamento industriais para otimizar os indicadores chave de desempenho (KPIs) de uma ou mais unidades da usina de processo, sendo que o método é **caracterizado** pelo fato de que compreende as etapas de:

receber, em um módulo de recebimento (108), uma pluralidade de dados de uma ou mais unidades de processamento industriais, em que a pluralidade de dados compreende características de matérias-primas, características de produtos intermediários, subprodutos e produtos finais, parâmetros de processo, parâmetros ambientais, demanda de mercado, disponibilidade de matérias-primas e condição de equipamentos do processo;

mesclar, em um módulo de fusão de nível de unidade (110), a pluralidade de dados recebida para obter conjunto de dados por unidade de cada uma dentre a uma ou mais unidades de processamento industriais, em que o conjunto de dados por unidade de cada unidade de processamento compreende uma frequência de amostragem desejada;

verificar, em um módulo de verificação (112), o conjunto de dados por unidade mesclado de uma ou mais unidades de processamento industriais, em que a presença de valores absurdos, porcentagem de disponibilidade, desvio padrão e

faixa interquartil de todas as variáveis da unidade de processamento são calculados;

pré-processar, em um módulo de pré-processamento de dados (114), a pluralidade de dados verificada para obter conjunto de dados pré-processado de cada uma dentre a uma ou mais unidades de processamento industriais, em que o pré-processamento é um processo iterativo que compreende as etapas de remoção de valor atípico, imputação de valores faltantes e agrupamento;

integrar, em um módulo de fusão de nível de empresa (116), a pré-processada de cada uma dentre a uma ou mais unidades de processamento industriais com um ou mais valores de variáveis simuladas de um ou mais modelos com base em física e uma ou mais entradas de domínio a partir do usuário para obter conjunto de dados de nível de empresa, em que o conjunto de dados por unidades é mesclado e sincronizado levando-se em conta os intervalos de tempo devido a tempos de permanência em diversas unidades, tempos de transporte de materiais entre uma ou mais unidades de processamento industriais e tempo de resposta de um ou mais sensores das unidades de processamento;

identificar, em um módulo de identificação de regime (118), um ou mais regimes de operação com o uso de uma ou mais técnicas de agrupamento no conjunto de dados de nível de empresa, em que uma ou mais técnicas de agrupamento

compreendem agrupamento com base em distância, agrupamento com base em densidade e agrupamento hierárquico;

determinar, em um módulo de estatística de linha de base (120), faixas de uma ou mais variáveis que correspondem aos KPIs do conjunto de dados de nível de empresa, com base em estatísticas de linha de base predefinidas e no um ou mais regimes de operação, em que as faixas determinadas de uma ou mais variáveis estão sendo usadas para gerar uma ou mais plotagens de KPIs durante o período de tempo em que a análise está sendo realizada;

selecionar, em um módulo de seleção de recurso (122), um ou mais recursos ou variáveis chave do conjunto de dados de nível de empresa para obter um superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa, em que a seleção de recurso é realizada em todos os conjuntos de dados por regime, bem como no conjunto de dados de nível de empresa;

desenvolver, em um módulo de construção de modelo (124), um ou mais modelos preditivos para cada KPI, em que o um ou mais modelos preditivos usam o conjunto de dados de nível de empresa e o superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa; e

otimizar, em um módulo de otimização (126), pelo menos um KPI com base em um ou mais modelos preditivos e restrições no um ou mais KPIs com o uso de uma ou mais técnicas de

otimização, em que uma ou mais técnicas de otimização incluem pesquisa de gradiente, programação linear, programação de objetivo, anelamento simulado e algoritmos evolucionários.

2. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que o pré-processamento é realizado em variáveis que têm uma porcentagem de disponibilidade predefinida e padrão de omissão predefinido.

3. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que a integração da pluralidade de dados pré-processada a partir de uma ou mais unidades industriais tem como base uma unidade de processo de linha de base predefinida.

4. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que a uma ou mais plotagens incluem plotagens de tendência de KPIs, plotagens de caixa dos KPIs, plotagens de dispersão e mapas de calor.

5. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que a seleção de recurso é realizada em dois estágios que compreendem:

no primeiro estágio, recursos importantes são obtidos a partir de uma ou mais técnicas de seleção de recurso, e

no segundo estágio, os recursos obtidos a partir do primeiro estágio são classificados com o uso do método de pontuação de média geométrica e combinados para obter um único superconjunto de um ou mais recursos.

6. Método, de acordo com a reivindicação 5, **caracterizado** pelo fato de que o recurso de pontuação menor dentre o um ou mais recursos selecionados no primeiro estágio é maior em relevância em relação a KPI.

7. Método, de acordo com a reivindicação 5, **caracterizado** pelo fato de que o primeiro estágio de uma ou mais técnicas de seleção de recurso compreende métodos com base em modelo e com base em não modelo.

8. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que a seleção de recursos é realizada em todos os conjuntos de dados por regime, bem como no conjunto de dados de nível de empresa.

9. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que as saídas para o usuário a partir do módulo de pré-processamento de dados (114) incluem uma lista de variáveis descartadas, número e porcentagem de valores atípicos removidos para cada variável, técnica usada para imputação de valores faltantes em cada variável, média, mediano e desvio padrão de cada variável antes e depois do pré-processamento e plotagens de tendência de todas as variáveis antes e depois do pré-processamento.

10. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que as saídas para o usuário a partir do módulo de fusão de nível de empresa (116) incluem uma lista de parâmetros simulados, e a faixa, média, mediano

e desvio padrão de todas as variáveis no conjunto de dados integrado.

11. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que as saídas para o usuário a partir do módulo de estatística de linha de base (120) incluem as faixas de variáveis que correspondem às faixas desejadas e indesejadas de KPIs, as faixas de KPIs em níveis de produtividade diferentes e os coeficientes de correlação entre KPIs e outras variáveis.

12. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que as saídas para o usuário a partir do módulo de estatística de linha de base (120) incluem as plotagens de tendência e plotagens de caixa de KPIs e outras variáveis, plotagens de dispersão entre KPIs e variáveis de interesse, e mapas de calor de valores médios dos KPIs.

13. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que as saídas para o usuário a partir do módulo de seleção de recurso (122) incluem o superconjunto de recursos e suas pontuações de importância para conjuntos de dados integrados e por regime, e plotagens de coordenadas paralelas dos recursos.

14. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que as saídas para o usuário a partir do módulo de discriminação e construção de modelo

incluem métricas de desempenho para todos os modelos preditivos, três principais modelos preditivos desenvolvidos com base no erro quadrático médio e erro médio absoluto, pontuações de robustez para os três principais modelos, pontuações de sensibilidade para todas as variáveis nos modelos robustos.

15. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que as saídas para o usuário a partir do módulo de discriminação e construção de modelo incluem também as plotagens de tendência dos valores reais e previstos do KPI, plotagens de dispersão de valores reais versus previstos do KPI e plotagens residuais de erro absoluto versus todas as variáveis nos modelos robustos.

16. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que as saídas para o usuário a partir do módulo de otimização (126) incluem os valores de variáveis que rendem KPIs ideais (pontos de operação de ótimo de Pareto) e valores ideais dos KPIs e plotagens de pontos de operação de ótimo de Pareto.

17. Sistema (100) para analisar uma pluralidade de dados a partir de uma ou mais unidades de processamento industriais para otimizar os indicadores chave de desempenho da indústria, sendo que o sistema é **caracterizado** pelo fato de que compreende:

uma memória (104) com instruções;

pelo menos um processador (102) acoplado de maneira comunicativa com a memória;

uma pluralidade de interfaces (106), em que a pluralidade de interfaces compreende interface gráfica de usuário, interface de servidor, uma interface de modelo com base em física e uma interface de resolvedor;

um módulo de recebimento (108) é configurado para receber uma pluralidade de dados de uma ou mais unidades de processamento industriais, em que a pluralidade de dados compreende características de matérias-primas, características de subprodutos, produtos finais e intermediários, parâmetros de processo e condição de equipamentos do processo;

um módulo de fusão de nível de unidade (110) é configurado para mesclar a pluralidade de dados recebida para obter conjunto de dados por unidade de cada uma dentre a uma ou mais unidades de processamento industriais, em que o conjunto de dados por unidade de cada unidade de processamento compreende uma frequência de amostragem desejada;

um módulo de verificação (112) é configurado para verificar o conjunto de dados por unidade mesclado da uma ou mais unidades de processamento industriais, em que a presença de valores indesejados, porcentagem de disponibilidade, desvio padrão e faixa interquartil de todas as variáveis da

unidade de processamento são calculados;

um módulo de pré-processamento de dados (114) é configurado para pré-processar a pluralidade de dados verificada para obter conjunto de dados pré-processado de cada uma dentre a uma ou mais unidades de processamento industriais, em que o pré-processamento é um processo iterativo que compreende as etapas de remoção de valor atípico, imputação de valores faltantes e agrupamento;

um módulo de fusão de nível de empresa (116) é configurado para integrar a pré-processada de cada uma dentre a uma ou mais unidades de processamento industriais com um ou mais valores de variáveis simuladas de um ou mais modelos com base em física e uma ou mais entradas de domínio do usuário para obter conjunto de dados de nível de empresa, em que o conjunto de dados por unidades é mesclado e sincronizado levando-se em conta os intervalos de tempo devido a tempos de permanência em diversas unidades, tempos de transporte entre uma ou mais unidades de processamento industriais e tempo de resposta de um ou mais sensores das unidades de processamento;

um módulo de identificação de regime (118) é configurado para identificar um ou mais regimes de operação com o uso de uma ou mais técnicas de agrupamento no conjunto de dados de nível de empresa, em que uma ou mais técnicas de agrupamento compreendem agrupamento com base em distância, agrupamento

com base em densidade e agrupamento hierárquico;

um módulo de estatística de linha de base (120) é configurado para determinar faixas de uma ou mais variáveis que correspondem aos KPIs do conjunto de dados de nível de empresa, com base em estatísticas de linha de base predefinidas e no um ou mais regimes de operação, em que as faixas determinadas de uma ou mais variáveis estão sendo usadas para gerar uma ou mais plotagens de KPIs durante o período de tempo em que a análise está sendo realizada;

um módulo de seleção de recurso (122) é configurado para selecionar um ou mais recursos do conjunto de dados de nível de empresa para obter um superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa, em que a seleção de recurso é realizada em todos os conjuntos de dados por regime, bem como no conjunto de dados de nível de empresa;

um módulo de construção de modelo (124) é configurado para considerar um ou mais modelos preditivos para cada KPI, em que o um ou mais modelos preditivos usam o conjunto de dados de nível de empresa e o superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa; e

um módulo de otimização (126) é configurado para otimizar pelo menos um KPI com base em um ou mais modelos preditivos e restrições no um ou mais KPIs com o uso de uma

ou mais técnicas de otimização, em que uma ou mais técnicas de otimização incluem pesquisa de gradiente, programação linear, anelamento simulado e algoritmos evolucionários.

18. Sistema, de acordo com a reivindicação 17, **caracterizado** pelo fato de que o um ou mais modelos com base em física são usados para o cálculo da uma ou mais variáveis simuladas.

19. Uma ou mais mídias de armazenamento de informações legíveis por máquina não transitórias **caracterizadas** pelo fato de que compreendem uma ou mais instruções que, quando executadas por um ou mais processadores de hardware, realizam ações que compreendem:

receber, em um módulo de recebimento (108), uma pluralidade de dados de uma ou mais unidades de processamento industriais, em que a pluralidade de dados compreende características de matérias-primas, características de produtos intermediários, subprodutos e produtos finais, parâmetros de processo, parâmetros ambientais, demanda de mercado, disponibilidade de matérias-primas e condição de equipamentos do processo;

mesclar, em um módulo de fusão de nível de unidade (110), a pluralidade de dados recebida para obter conjunto de dados por unidade de cada uma dentre a uma ou mais unidades de processamento industriais, em que o conjunto de dados por unidade de cada unidade de processamento compreende uma

frequência de amostragem desejada;

verificar, em um módulo de verificação (112), o conjunto de dados por unidade mesclado de uma ou mais unidades de processamento industriais, em que a presença de valores absurdos, porcentagem de disponibilidade, desvio padrão e faixa interquartil de todas as variáveis da unidade de processamento são calculados;

pré-processar, em um módulo de pré-processamento de dados (114), a pluralidade de dados verificada para obter conjunto de dados pré-processado de cada uma dentre a uma ou mais unidades de processamento industriais, em que o pré-processamento é um processo iterativo que compreende as etapas de remoção de valor atípico, imputação de valores faltantes e agrupamento;

integrar, em um módulo de fusão de nível de empresa (116), a pré-processada de cada uma dentre a uma ou mais unidades de processamento industriais com um ou mais valores de variáveis simuladas de um ou mais modelos com base em física e uma ou mais entradas de domínio a partir do usuário para obter conjunto de dados de nível de empresa, em que o conjunto de dados por unidades é mesclado e sincronizado levando-se em conta os intervalos de tempo devido a tempos de permanência em diversas unidades, tempos de transporte de materiais entre uma ou mais unidades de processamento industriais e tempo de resposta de um ou mais sensores das

unidades de processamento;

identificar, em um módulo de identificação de regime (118), um ou mais regimes de operação com o uso de uma ou mais técnicas de agrupamento no conjunto de dados de nível de empresa, em que uma ou mais técnicas de agrupamento compreendem agrupamento com base em distância, agrupamento com base em densidade e agrupamento hierárquico;

determinar, em um módulo de estatística de linha de base (120), faixas de uma ou mais variáveis que correspondem aos KPIs do conjunto de dados de nível de empresa, com base em estatísticas de linha de base predefinidas e no um ou mais regimes de operação, em que as faixas determinadas de uma ou mais variáveis estão sendo usadas para gerar uma ou mais plotagens de KPIs durante o período de tempo em que a análise está sendo realizada;

selecionar, em um módulo de seleção de recurso (122), um ou mais recursos ou variáveis chave do conjunto de dados de nível de empresa para obter um superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa, em que a seleção de recurso é realizada em todos os conjuntos de dados por regime, bem como no conjunto de dados de nível de empresa;

desenvolver, em um módulo de construção de modelo (124), um ou mais modelos preditivos para cada KPI, em que o um ou mais modelos preditivos usam o conjunto de dados de nível de

empresa e o superconjunto de um ou mais recursos selecionados do conjunto de dados de nível de empresa; e

otimizar, em um módulo de otimização (126), pelo menos um KPI com base em um ou mais modelos preditivos e restrições no um ou mais KPIs com o uso de uma ou mais técnicas de otimização, em que uma ou mais técnicas de otimização incluem pesquisa de gradiente, programação linear, programação de objetivo, anelamento simulado e algoritmos evolucionários.

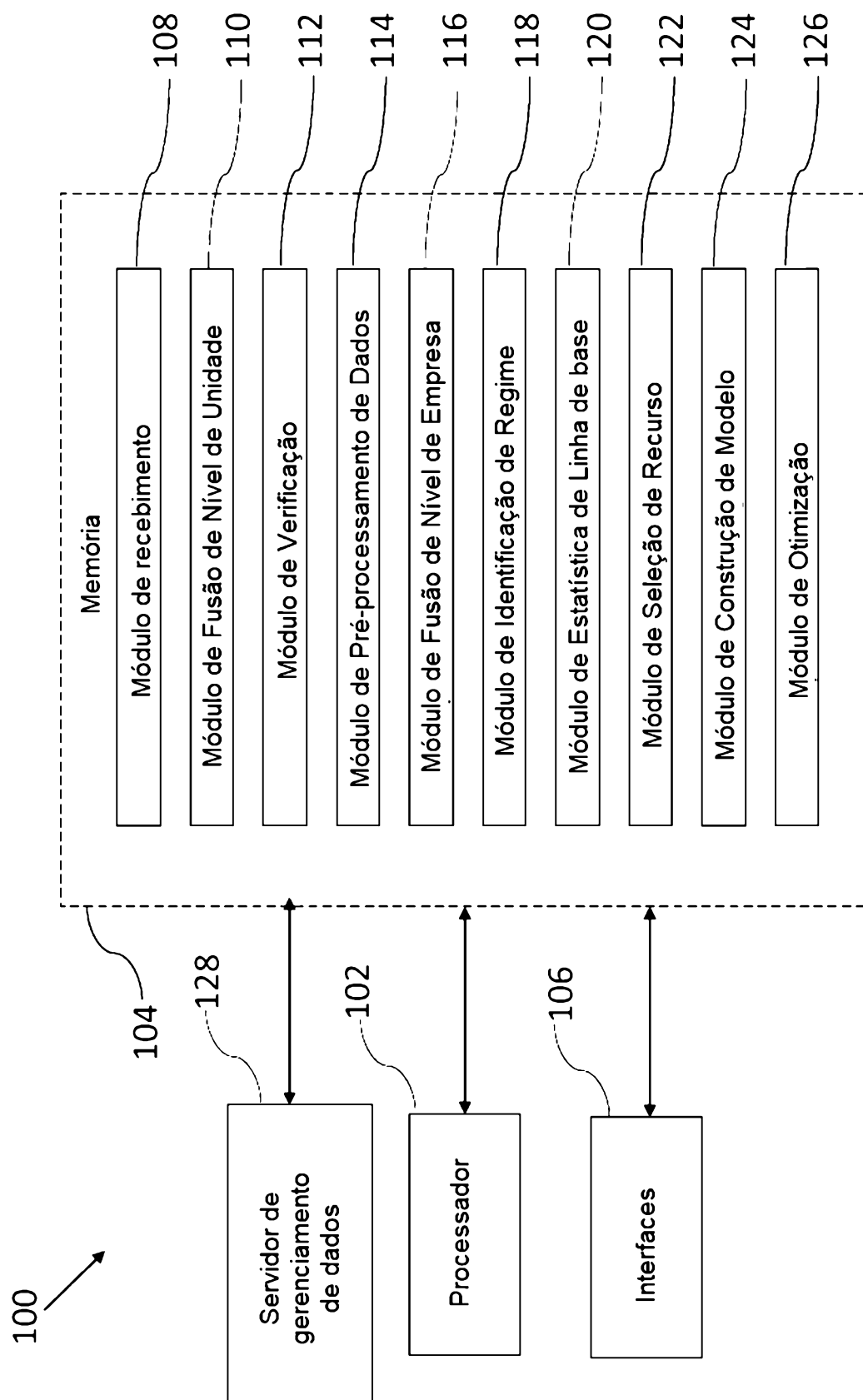


Figura 1

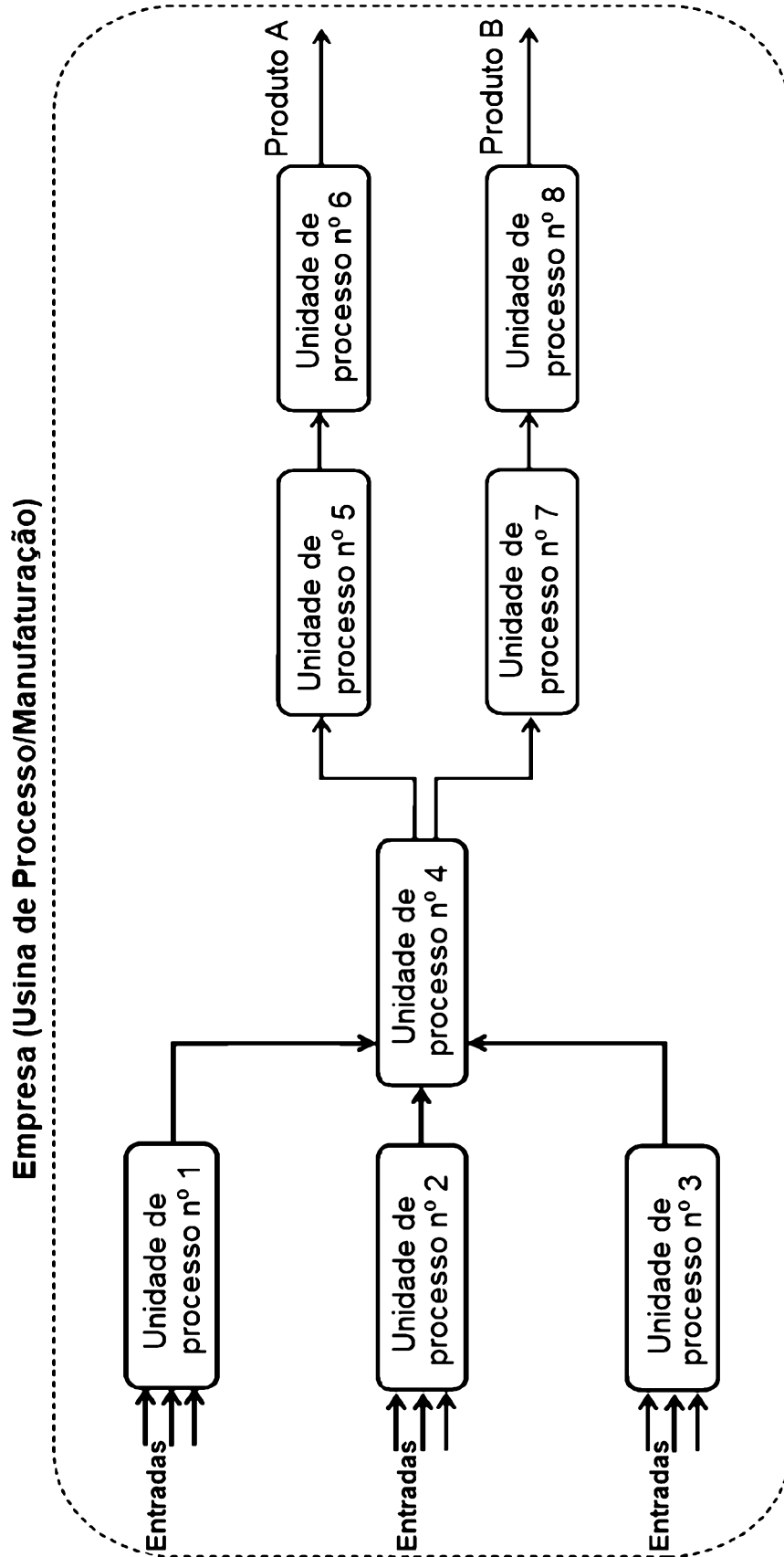


Figura 2

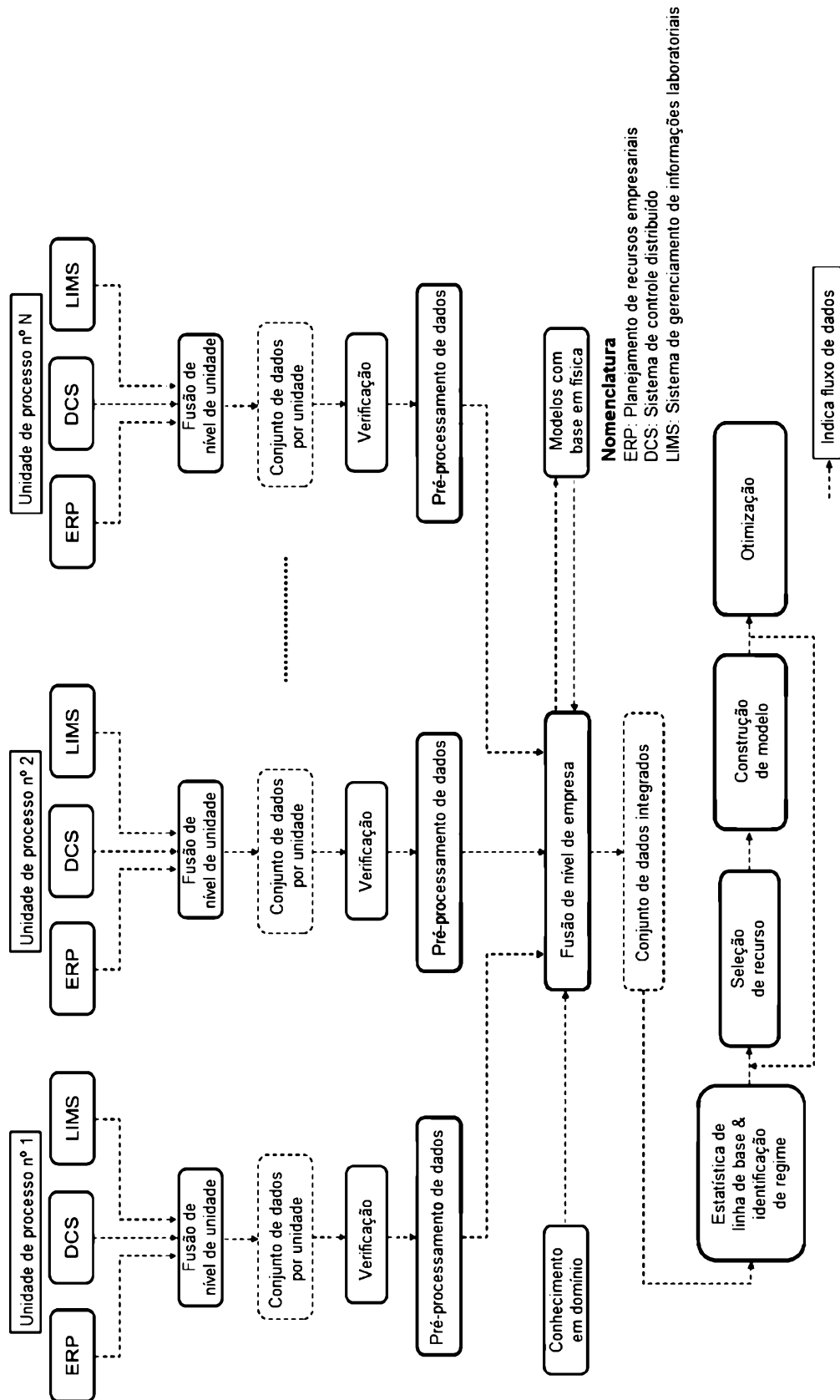


Figura 3

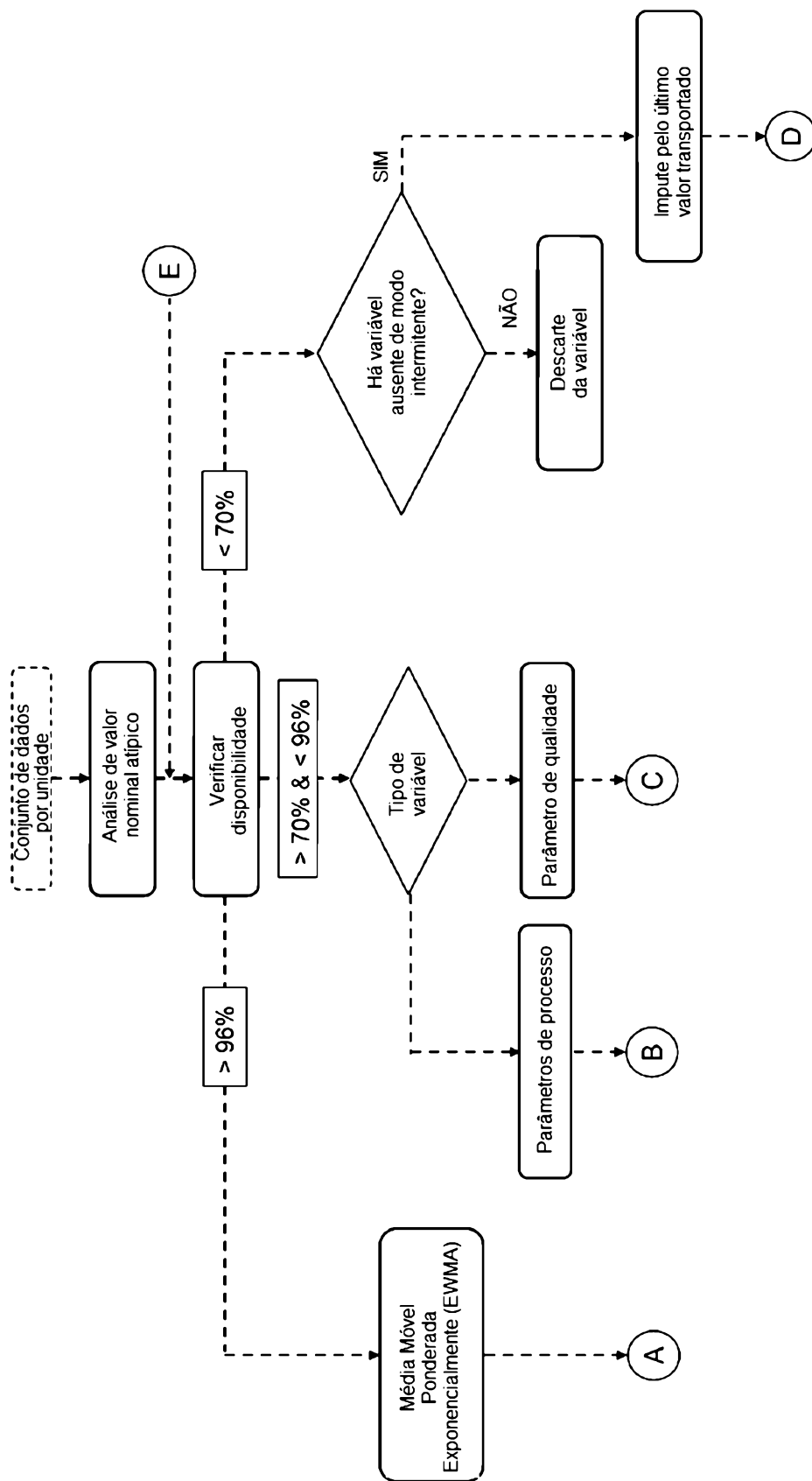


Figura 4(a)

Requerente: Tata Consultancy Services Limited

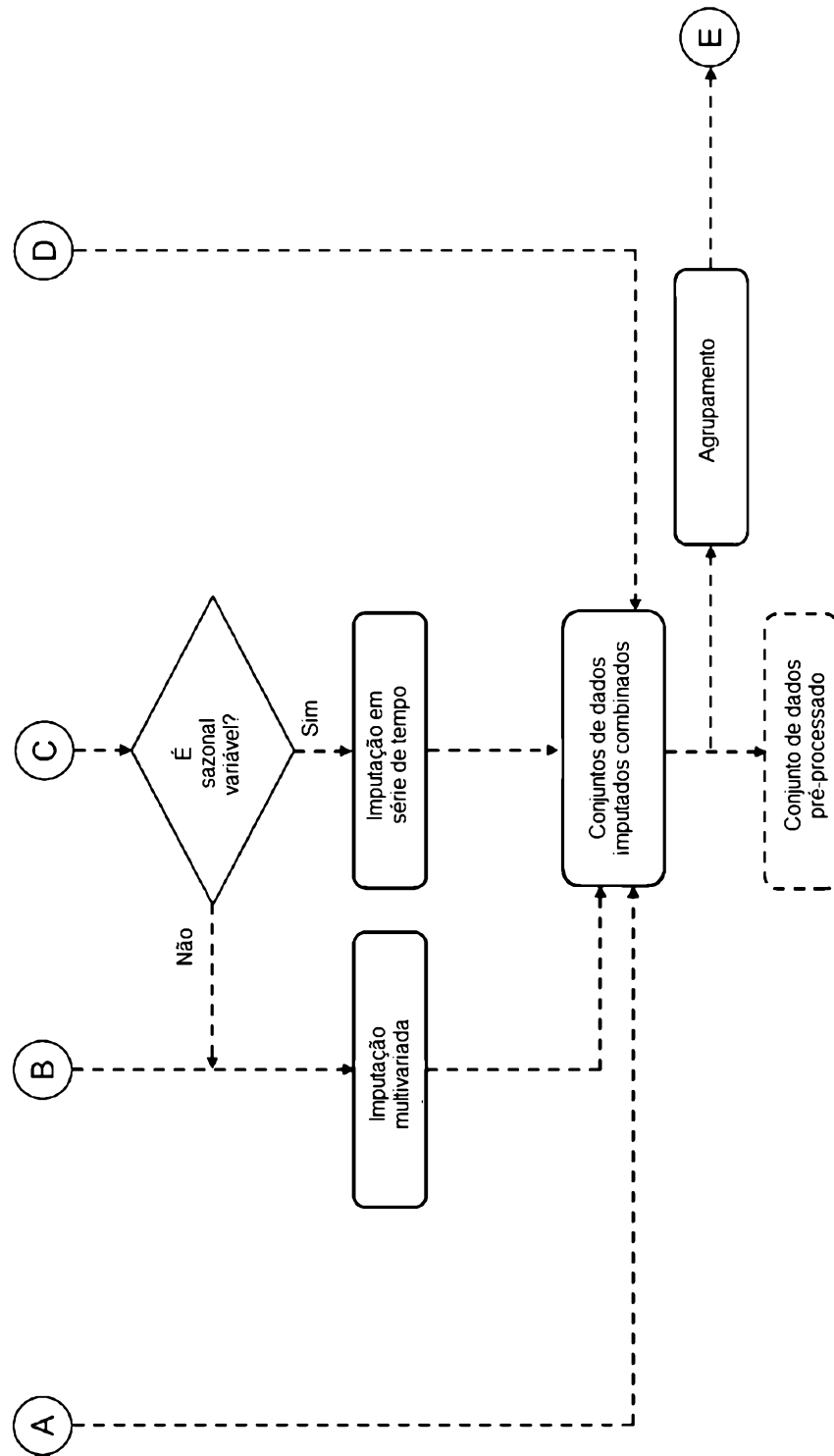


Figura 4(b)

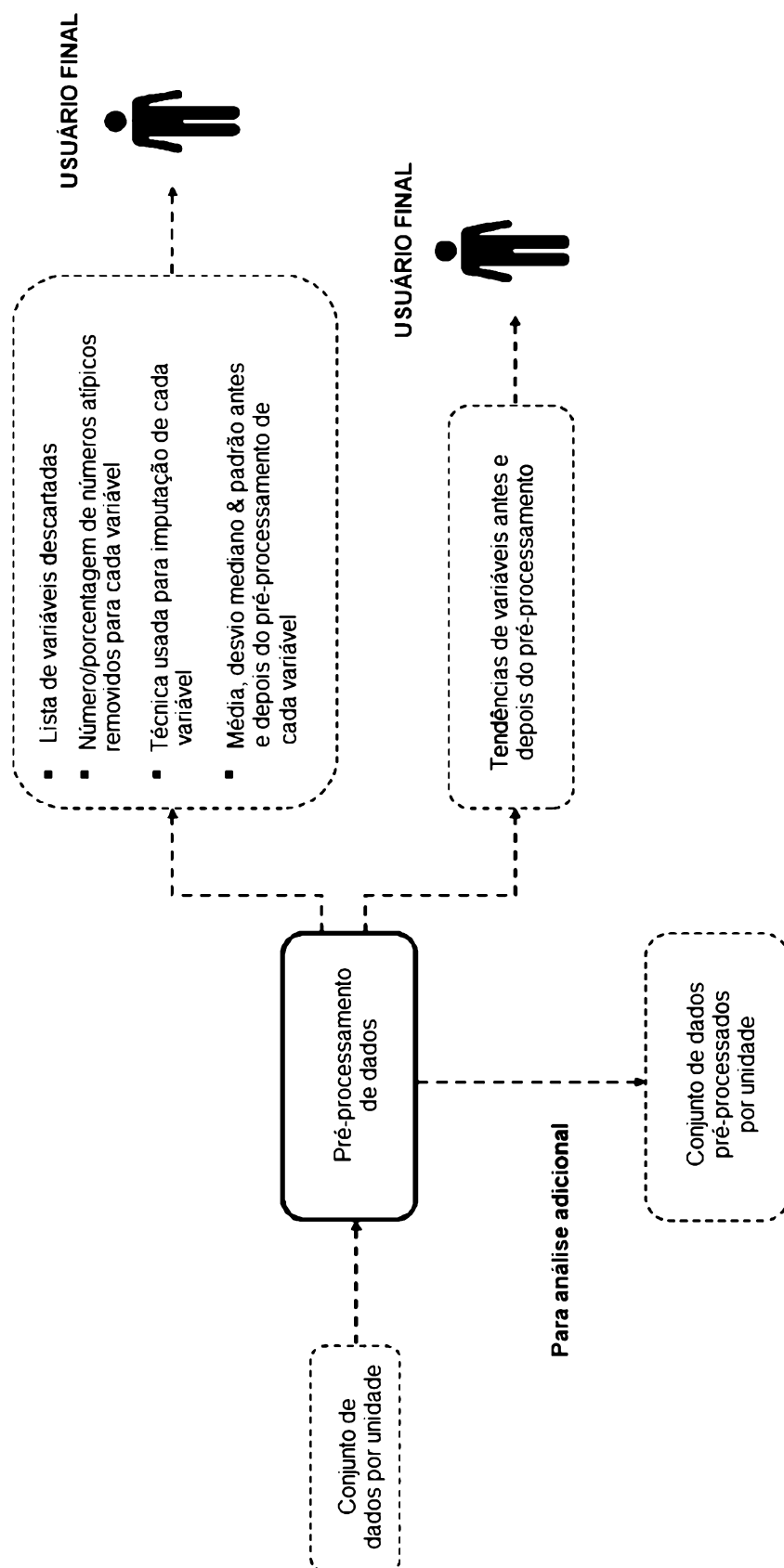


Figura 5

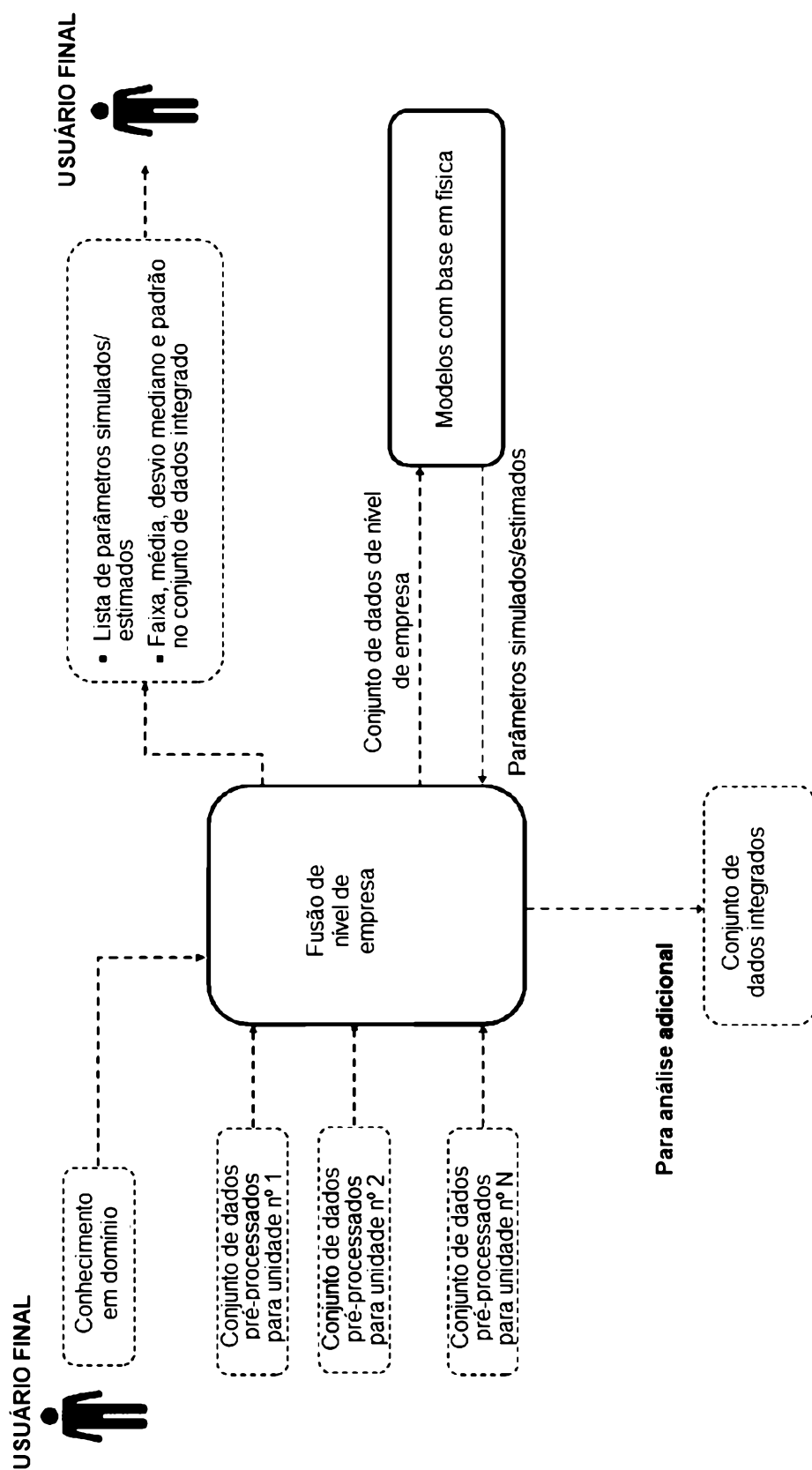


Figura 6

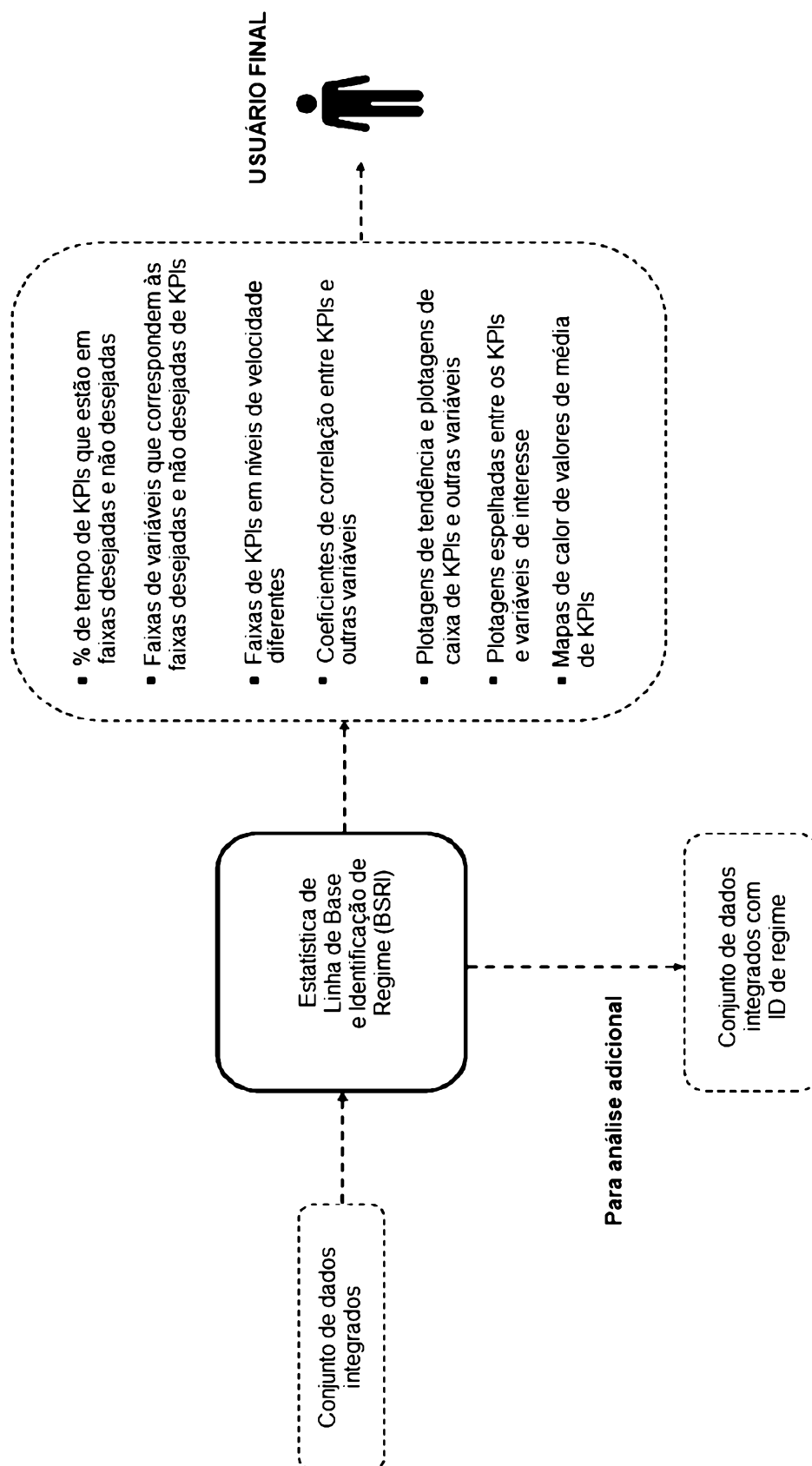


Figura 7

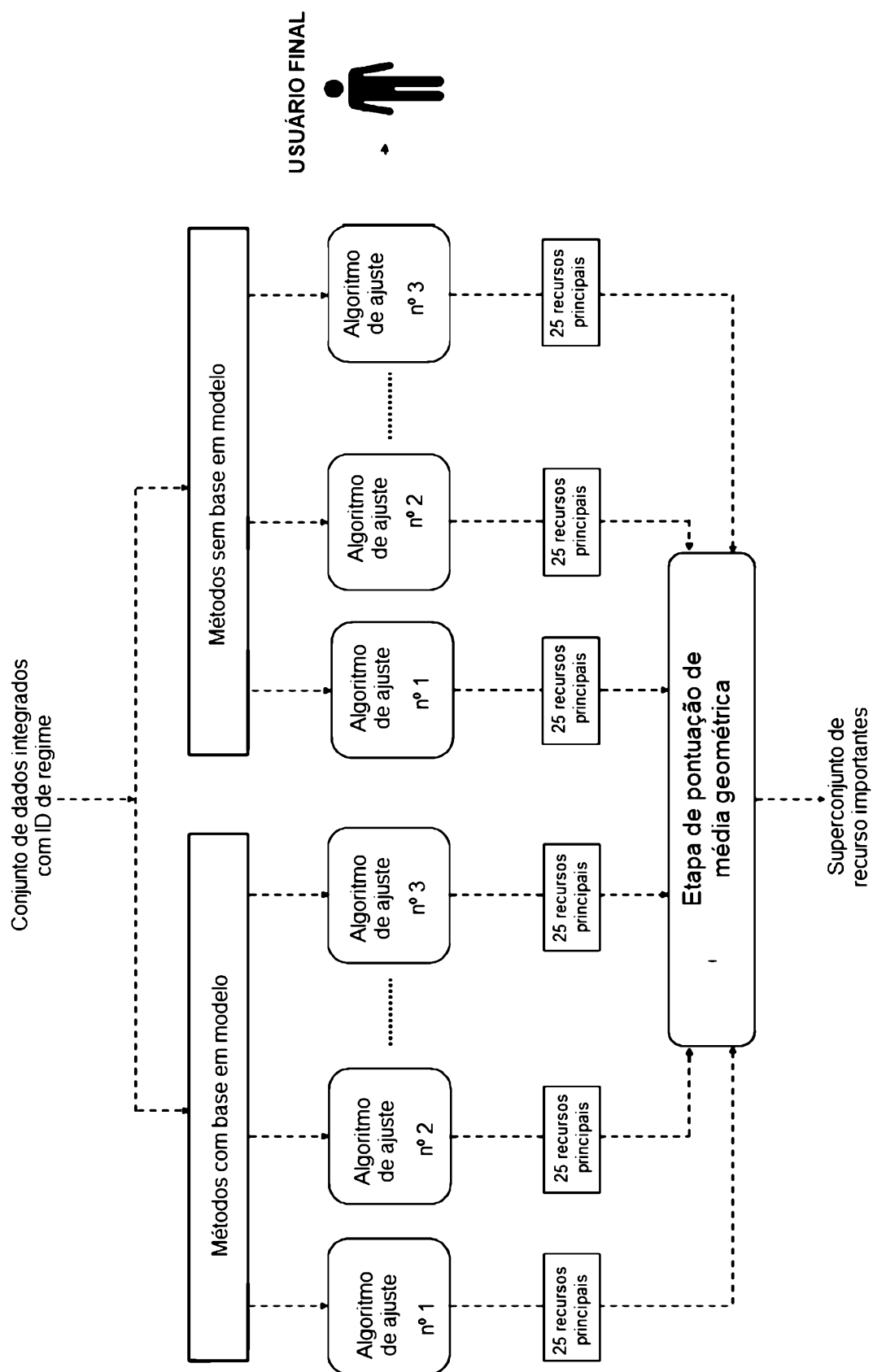


Figura 8

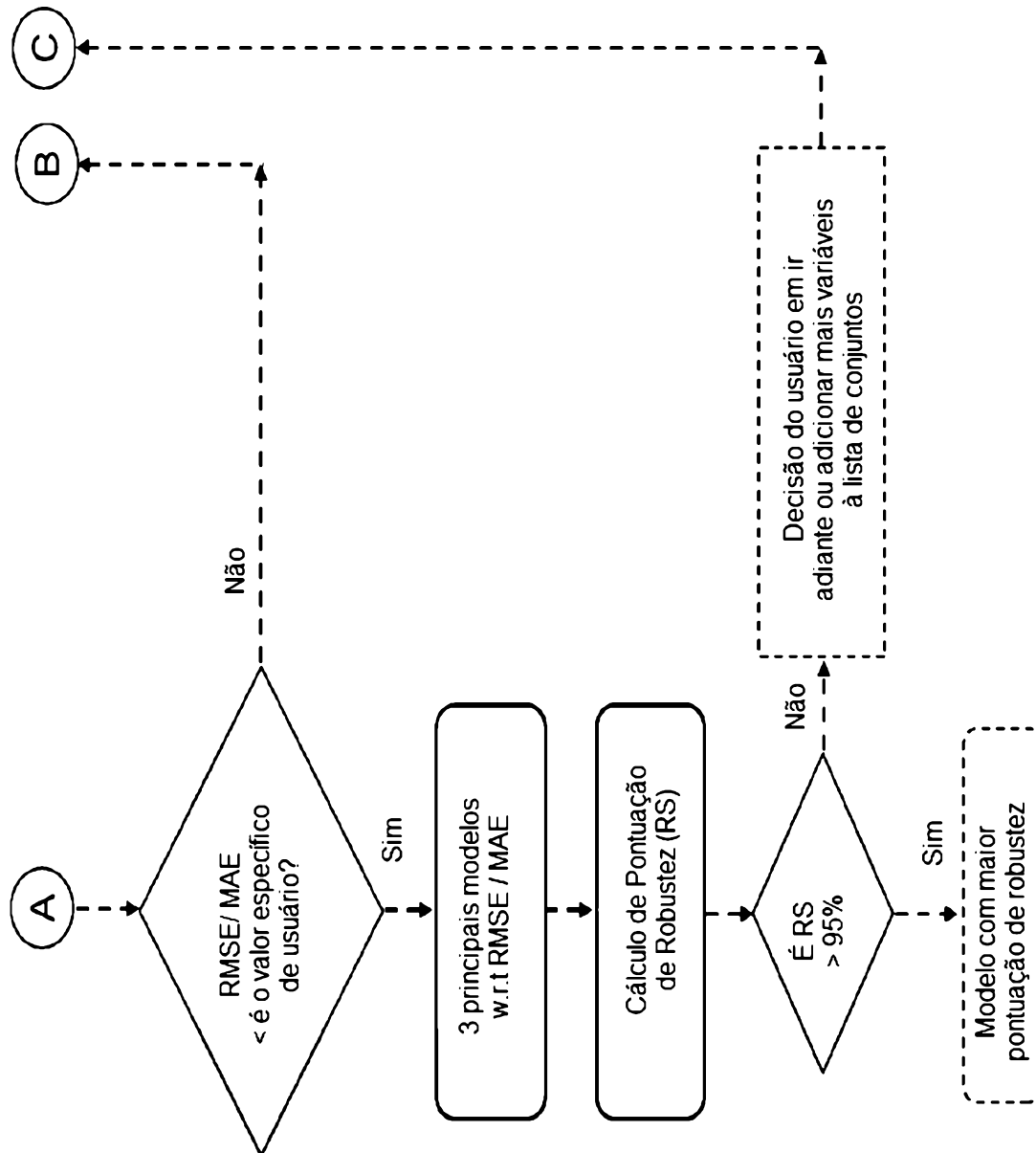


Figura 9(b)

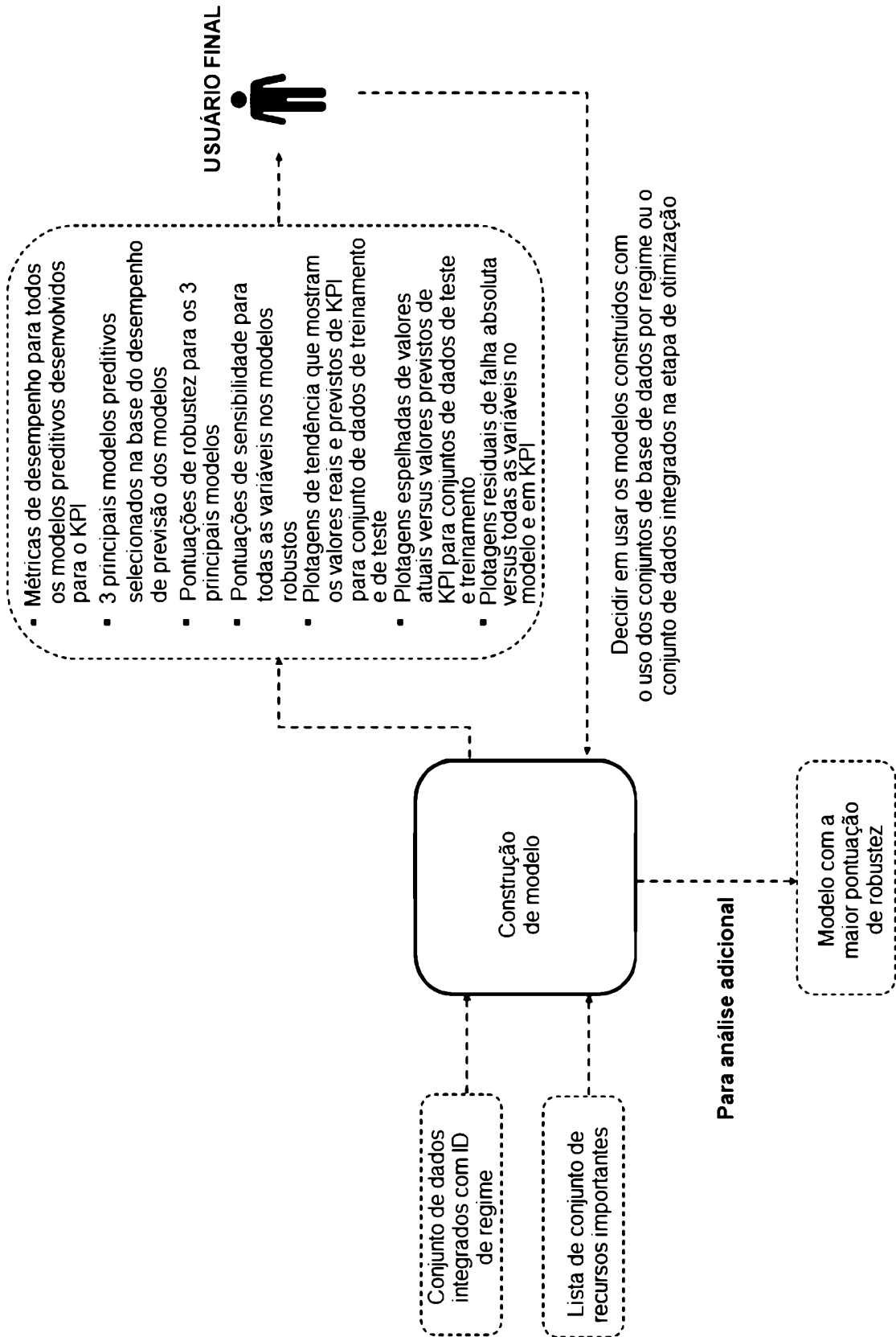


Figura 10

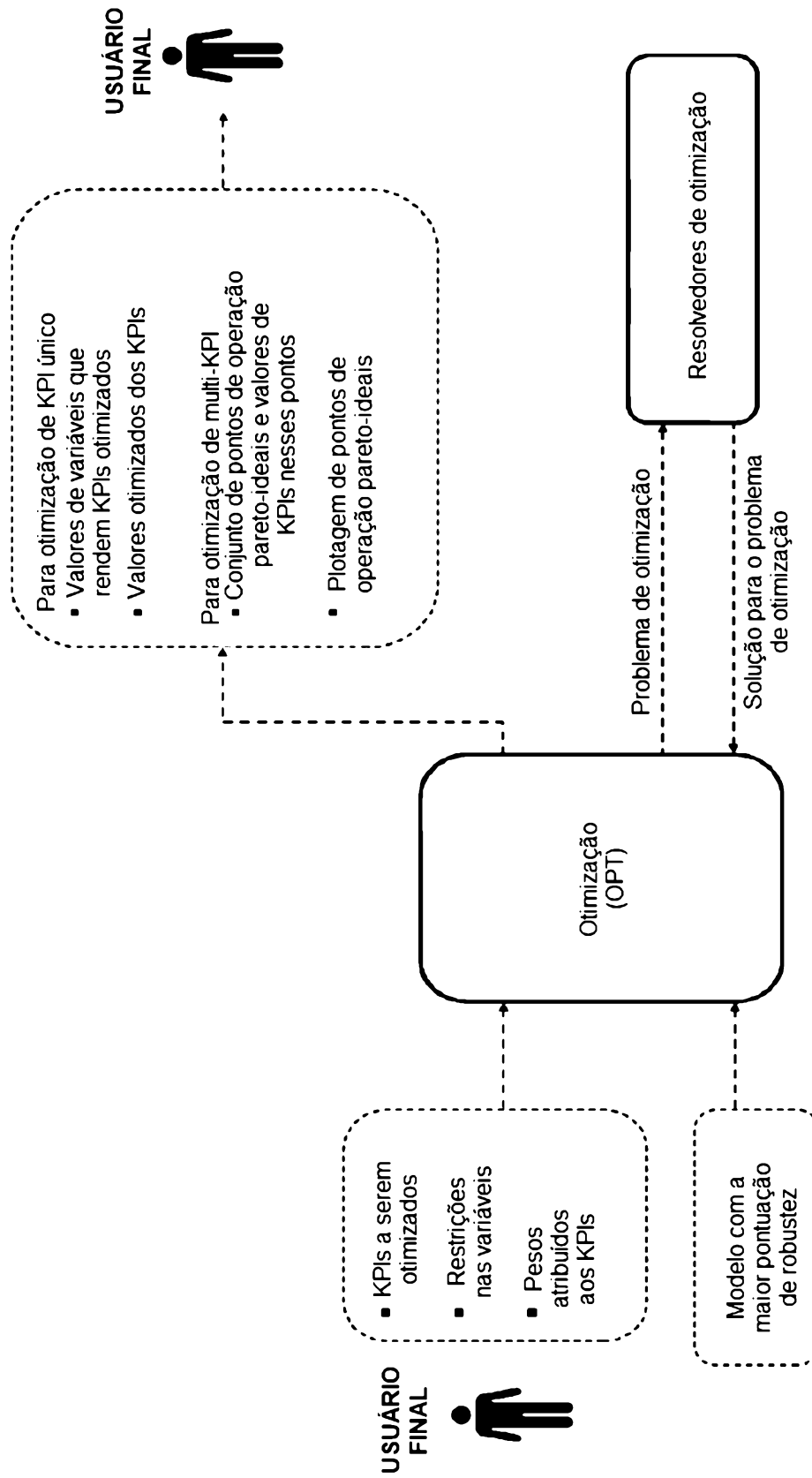


Figura 11

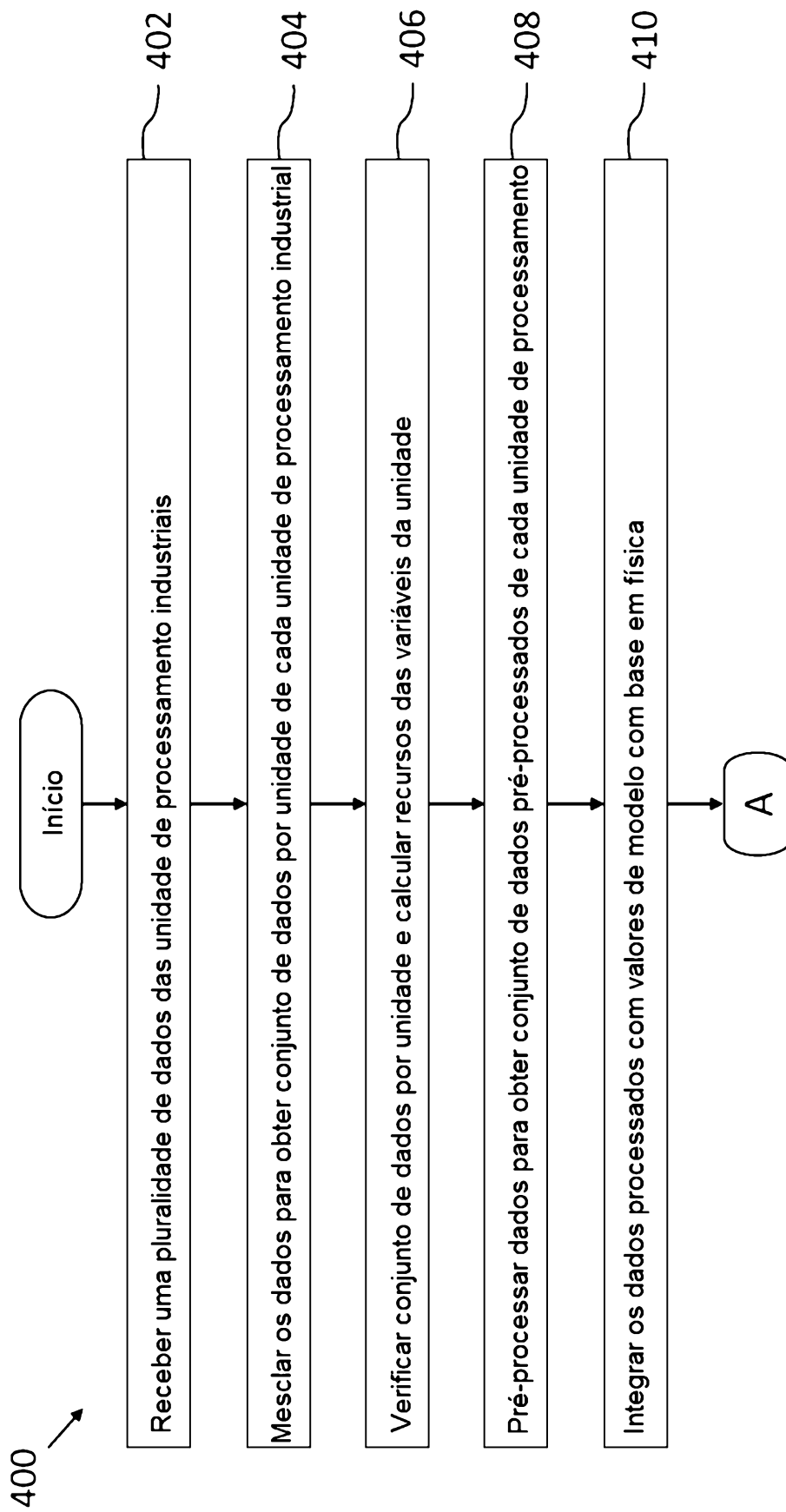


Figura 12(a)

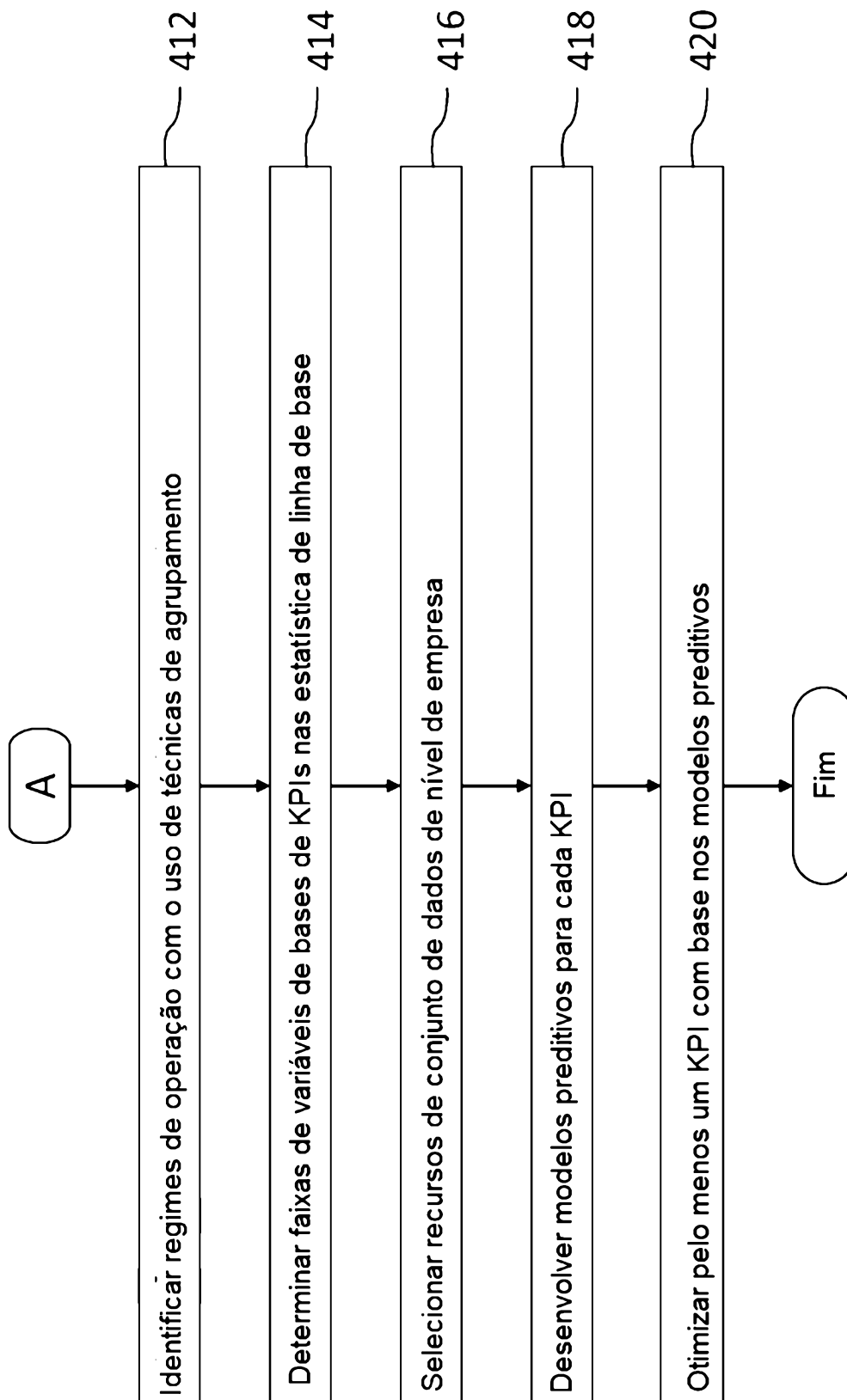


Figura 12(b)

RESUMO**"MÉTODO E SISTEMA PARA OTIMIZAÇÃO COM BASE EM DADOS DE INDICADORES DE DESEMPENHO EM INDÚSTRIAS DE FABRICAÇÃO E PROCESSO"**

Trata-se de um sistema e método para realizar a otimização com base em dados de indicadores de desempenho de usinas de fabricação e processo. O sistema consiste em módulos para coletar e mesclar dados a partir de unidades de processamento industriais, pré-processando os dados para remover valores atípicos e omissão. Adicionalmente, o sistema gera saídas personalizadas a partir de dados e identifica variáveis importantes que afetam um determinado indicador de desempenho de processo. O sistema também constrói modelos preditivos para indicadores chave de desempenho que compreendem os recursos importantes e determina pontos de operação para otimizar os indicadores chave de desempenho com intervenção de usuário mínima. Em particular, o sistema recebe entradas a partir de usuários nos indicadores chave de desempenho a serem otimizados e notifica os usuários de saídas a partir de diversas etapas na análise que ajudam os usuários a gerenciar de modo eficaz a análise e tomar decisões operacionais adequadas.