This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3014620, IEEE Access

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2020 Doi Number

Finding Misstatement Accounts in Financial Statements through Ontology Reasoning

LIMING CHEN[1], BAOXIN XIU[2], ZHAOYUN DING[1]

[1]Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha 410073 China
[2]School of Systems Science and Engineering, Sun Yat-sen University, Guangzhou 510006 China

Corresponding author: Baoxin Xiu (e-mail: baoxinxiu@163.com).

ABSTRACT Finding misstatement accounts in financial statements, is a key problem of fraud detection. Potential applications include external audit, internal controls, investment decision and securities market regulation. However, most existing intelligent methods just detect financial statements fraud at the company level, while little research can detect financial statements fraud at the account level. For this, to achieve intelligent fraud detection at the accounts level, an ontology-based fraud detection framework was proposed. To be specific, the proposed framework mainly combines the articulation between different accounts and periods, and 30 financial indicators (ratios) as the knowledge basis of ontology. Notably, with OWL (*Ontology Web Language*), SWRL (*Semantic Web Rule Language*) and **Protégé** ontology editor, the case study not only completed the fraud detection in a fast and timely manner, but also provided logical explanation and risk warning at the accounts level. This fully shows the great advantages and applicability of the proposed framework in the detection of misstatements accounts. Moreover, the proposed framework is of great significance for timely detection, prevention and response of financial statements fraud. More importantly, the proposed framework opens-up a new direction of using ontology reasoning techniques to find misstatement accounts in financial statements, which provides an interpretable and fine-grained way for fraud detection.

INDEX TERMS Misstatement accounts, fraud detection, financial indicators (ratios), ontology reasoning

I. INTRODUCTION

On April 2, 2020, *Luckin Coffee* released a notice admitting that it had made 2.2 billion *yuan* of false transactions, which has not only seriously damaged the interests of investors, creditors and other stakeholders, but also brought serious harm to social and economic life. There is no denying that traditional financial fraud prevention methods, such as internal control [1, 2] and external audit, have made some progress. However, from the world-beating *Enron Incident* [2] to recent *Luckin Coffee* fraud, financial statements fraud never seems to stop, which confuses auditors, stakeholders and regulators all the time. That is, for investors, regulators, creditors, auditors as well as other stakeholders, timely and effective financial fraud risk early warning framework is urgently needed to be put forward.

Considering the recent advances of artificial intelligence in recent years, intelligent financial statements fraud detection may provide a new solution. On the one hand, the data-driven methods, such as machine learning, have made great progress in financial statements fraud detection. On the other hand, in many cases, financial statements detection with machine learning might encounter three bottlenecks: insufficient data, imbalanced data distribution and poor interpretability. In fact, apart from machine learning [3], knowledge engineering [4] is also an effective method of artificial intelligence. Moreover, knowledge engineering, especially ontology model, not only has strong knowledge representation ability and reasoning ability, but also can provide adequate explanation, which is very crucial for financial statements fraud detection and risk warning [5]. Therefore, intelligent ontology reasoning may be a powerful tool for financial statements fraud detection. Of course, some people have done the research of financial statements fraud detection based on ontology [6]. However, the existing intelligent detection methods, whether based on machine learning, data mining [7], or ontology model [6], both are company level detection. That is to say, these methods can only give a conclusion whether the company is fraudulent or not, but cannot focus on specific risk accounts, or financial

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3014620, IEEE Access

statements items. For this, this paper tries to propose an ontology-based financial fraud detection framework to find misstatement accounts in financial statements.

To find misstatement accounts in financial statements, a fundamental analysis of financial statements fraud is essential. In essence, the vast majority of financial statements fraud is to increase profits [8, 9]. As a result, financial statements fraud is mainly reflected in the key amounts fraud, or key financial statements items fraud. In other words, to some extent, the financial fraud detection problem, can be defined as the detection of misstatement accounts. For this, the items and articulation in financial statements become the major concern of the proposed framework. Based on this observation, the proposed framework takes the key raw financial indicators from financial statements as input, as well as deduces the derivative financial ratios and reasoning results. Among them, the introduce of derivative financial ratios makes the misstatement accounts detection simple and easy. In particular, the difference between the growth of *inventory* and *operating revenue* was innovatively designed to detect fraudulent *operating revenue*. Similarly, other derivative financial ratios, such as *cash ratio of net profit*, *return on cash* and *non-operating revenue ratio*, all can play a good warning role in finding various misstatement accounts.

In terms of the process, the ontology-based misstatement accounts detection framework mainly involves three steps. Considering the root of profit fraud, the first step is to select the possible fraud accounts from three financial statements (*Balance Sheet*, *Income Statement* and *Cash Flow Statement*) and prepare the essential domain knowledge (the articulation in financial statements) for ontology construction. The second step is to formalize the prepared knowledge and carry out consistency check with OWL (*Ontology Web Language*), SWRL (*Semantic Web Rule Language*) [10] and **Pellet** infer engine. Lastly, with the help of the **Pellet** inference engine, the reasoning results of misstatement accounts detection can be generated and used for risk warning for stakeholders, auditors and regulators. Notably, the proposed framework mainly includes five contributions: 1) accounts level fraud detection for the first time; 2) interpretable and timely fraud warning; 3) combination of time dimension and statements dimension; 4) collaborative filtering of inventory and operating revenue; 5) flexible adjustment mechanism based on thresholds.

For illustration, the rest of this paper is structured as follows: Section 2 reviews some related work of ontology-based decision support and intelligent financial statements fraud detection; Section 3 briefly illustrates the ontology-based misstatement accounts detection process in our framework; Section 4 introduces the construction of the ontology, ontology evolutional mechanisms, and rule-based ontology reasoning method; at last, Section 5 justifies the proposed ontological approaches through a case study in retail industry (*Luckin Coffee*) and Section 6 concludes the paper and presents some future work.

II. RELATED WORK

As far as we know, the current financial statements fraud detection based on ontology model is mainly done by Xiao-Bo, Guang-Chao et al. (2018) [6]. To be brief, Xiao-Bo, Guang-Chao et al. (2018) [6] developed a financial statements fraud detection system, which is based on a financial statement detection ontology and rules extracted from a C4.5 decision tree algorithm. That is, there are few researches in ontology-based financial statements fraud detection. Most relevant studies are concentrated on ontology-based decision support and intelligent financial statements fraud detection. For this, the following related work mainly focuses on ontology-based decision support and intelligent financial statements fraud detection.

A. ONTOLOGY-BASED DECISION SUPPORT

Ontology-based decision support, refers to providing decision support through ontology model. Roughly speaking, ontology-based decision support mainly involves three aspects: financial applications, clinical support and intelligent manufacturing.

1) FINANCIAL APPLICATIONS

In ontology-based financial applications, financial risk management accounts for the largest proportion. More specifically, ontology-based financial management mainly includes bankruptcy prediction [11], catastrophic financial systems risk control [12], financial statements fraud detection [6], insurance company activity modeling [13] and other applications [14, 15]. To prevent systemic financial risks, Organ and Stapleton (2017) [12] focused on the financial crisis from 2007 to 2009 and proposed to incorporate human factors into the ontology model. For the risk of company bankruptcy, an effective ontology model based on financial statements was proposed by Martin, Manjula et al. (2011) [11]. which combined financial domain ontology model with association rule mining algorithm and *Z-score* model. Furthermore, to detect financial statements fraud, Xiao-Bo, Guang-Chao et al. (2018) [6] presented a knowledge-based financial statements fraud detection system, which introduced a machine-learning algorithm to discover the financial variables and fraud detection rules, and leveraged an ontology with inference engine to discover implicit knowledge. Apart from risk management, ontology-based financial applications also involves opinion mining of financial news [16], which proposed an ontology-driven approach to semantically describe relations between concepts in the financial news domain.

To sum up, a variety of financial applications, including financial statements fraud detection, have leveraged ontology for decision support. However, these studies are hardly specific to the account level, which is inconvenient for decision making.

2) CLINICAL SUPPORT

For a long time, ontology-based clinical support has always been an unavoidable topic [4]. Generally speaking, ontology-

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

based disease diagnosis mainly refers to providing decision support for the disease diagnosis and treatment with ontology model [17].

Taking one step further, diagnostic support with ontology, mainly includes the diagnosis of diabetes [18], infectious disease [19] and generic multi-agent diagnostic systems [20]. For example, a semantically interpretable FRBS (*Fuzzy Rule-Based System*) framework [18] was proposed and implemented for diagnosis of diabetes. Moreover, its empirical results on a real dataset can indicate how the proposed system helps physicians and patients to accurately diagnose diabetes mellitus.

In disease treatment, ontology-based decision support mainly focuses on cancer treatment [21] and antibiotic treatment [17, 22, 19]. Notably, Shen, Colloc et al. (2018) [21] proposed a decision support system called DSS (*Decision Support System*). Specifically, DSS mainly used CBR (*Case-Based Reasoning*) to consider disease manifestations and provided physicians with treatment solutions from similar previous cases for reference. Experimentally, the ontology-based DSS obtained 84.63% accuracy [21] in disease classification.

3) INTELLIGENT MANUFACTURING

Ontology-based decision support systems have also been used to manufacturing companies. To the best of our knowledge, these studies are mainly devoted to solving the problems of industrial chain resilience [23, 24], production system reconfiguration [25, 26] and production process selection in manufacturing [27, 28, 24]. Firstly, to address the uncertainly challenge in manufacturing and supply chain, a rule-based ontology model [23] was constructed to enhance supply chain resilience. Furthermore, in order to optimize the manufacturing process, Mabkhot, Amri et al. (2020) [25] designed a knowledge-based multi-criteria decision support system to suggest candidate configurations and select a suitable configuration of manufacturing systems. Additionally, a rule-based ontology reasoning method [24] was proposed to provide decision support for steel manufactures as well.

According to the related work above, the interpretability and reasoning of ontology are naturally suitable for intelligent decision support, especially for financial statements fraud detection. Inspired by this observation, this paper finally adopts ontology reasoning to find misstatement accounts in financial statements.

B. INTELLIGEN FINANCIAL STATEMENTS FRAUD DETECTION

Data and methods together determine the performance of financial statements fraud detection. In light of this, the following will give a review of relevant studies from data dimension and methods dimension.

1) DATA DIMENSION

Considering the close relationship between financial fraud detection and financial forecast, the review of data dimension will combine the two aspects together. Data, used for financial statements fraud detection and financial forecasting, mainly involves two types: textual data [29] and technical indicators [30]. The related research mainly takes textual data or technical indicators data as input and modeling as the financial time series forecasting problem [31, 32]. Taking one step further, textual data and technical indicators will be introduced separately as follows.

To utilize textual data for financial predicting, natural language based financial forecasting (NLFF) has been a research field. To be specific, financial reports [29], social media [33] and news [34] are most commonly adopted [29]. Generally speaking, most existing researches tend to analyze these text data by natural language processing techniques, especially sentiment analysis. The main idea is to take text data as input and output stock market trend or stock prices prediction [33]. Based on this, an immediate response to market sentiment can be achieved. For example, Chiong, Fan et al. (2018) [35] have considered financial news data for financial market prediction and achieved superior experimental results. Besides, technical indicators are generally related to stock prices or stock prices movement prediction. Concretely, relevant technical indicators mainly involves *Opening Price*, *Closing Price*, *Highest Price*, *Lowest Price*, *Trading Volume*, *Turnover Rate* [36, 37] and their derivative indicators [30, 38]. For example, Relative Strength Index was adopted to predict significant stock price changes and got superior experimental results by Kamalov (2019) [39]. In addition, multi-source data was also considered for predicting stock prices [31, 5, 40]. For example, Zhou, Gao et al. (2020) [40] employed multiple heterogeneous data sources, including historical transaction data, technical indicators, stock posts, news and Baidu index, to predict the directions of stock price movements.

On the one hand, to some extent, the method based on textual data and technical indicators, have achieved good results in financial forecasting, especially in timely risk warning and investment decision-making. On the other hand, these methods usually lack the analysis of company fundamentals, which may lead to uncontrolled risks. In addition, both textual data and technical indicators are likely to suffer from imbalance and high computational complexity. For the above problems, the proposed framework focuses on the basic analysis of companies, and chooses three major financial statements and their derivative indicators as the input of ontology model.

2) METHODS DIMENSION

In terms of quantity, the researches of intelligent financial statements fraud detection usually tend to adopt data mining [41-43], or machine learning [3, 44] methods. Besides, some researches have also attempted to use genetic algorithm [45], evolutionary algorithm [46] and other methods [47, 6] to detect financial statements fraud. In detail, major intelligent methods used for financial statements fraud detection are shown in Table I. Undeniably, these methods have achieved remarkable results in a certain range. However, these methods

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

are mainly dedicated to identifying whether a company has the problem of financial statements fraud. That is to say, the existing researches are almost the financial statements fraud detection at the company level. In other words, for financial statements fraud detection, most existing intelligent methods are hard to identify the specific fraud accounts or accounting subjects and lack of interpretability.

TABLE I
INTELLIGENT METHODS USED FOR FINANCIAL STATEMENTS FRAUD DETECTION

	Methods	Fields of Methods	References
1	Meta-Learning	Machine Learning	(Abbasi, Albrecht et al. 2012) [48]
2	Multiple Instance Learning (MIBoost, miGraph and CKN)	Machine Learning	(Tang, Peng et al. 2016) [49]
3	Multi-Layer Feed Forward Neural Network(MFFNN), Probabilistic Neural Network (PNN), Support Vector Machine (SVM), Multinomial Log-Linear Model (MLM), and Discriminant Analysis (DA)	Machine Learning	(Omidi, Min et al. 2019) [50]
4	Ensemble Learning (IMF, a novel combiner method classification algorithm)	Machine Learning (Data Mining)	(Perols 2008) [51]
5	Ensemble Learning (RUSBoost [52])	Machine Learning (Data Mining)	(Seiffert, Khoshgoftaar et al. 2010) [52]
6	Ensemble Methods, Bayesian Belief Networks	Machine Learning (Data Mining)	(Hajek and Henriques 2017) [53]
7	Decision Trees, Neural Networks and Bayesian Belief Networks	Machine Learning (Data Mining)	(Dan 2017) [7]
8	Logistic Regression, Support Vector Machine, Artificial Neural Network, Bagging, Decision Tree (C4.5) and Stacking	Machine Learning (Data Mining)	(Perols 2011) [44]
9	Queen Genetic Algorithm-Support Vector Machine (QGA-SVM), Decision Tree (C4.5), Logistic Regression, Back-Propagation Neural network (BPN), K-Nearest Neighbors (KNN), genetic algorithms-support vector machine (GA-SVM), and particle swarm optimization-support vector machine (PSO-SVM)	Machine Learning (Data Mining)	(Chen, Liou et al. 2019) [54]
10	Logistic Regression (LR), Back-propagation neural network (BPNN), Decision tree (DT, C5.0), Support vector machine (SVM) and a Hybrid Classifier	Machine Learning (Data Mining)	(Xinping and Yan 2012) [55]
11	The genetic algorithm considers combinations of variables and the interactions of variables across time when generating candidate patterns.	Genetic Algorithms	(Hoogs, Kiehl et al. 2007) [45]
12	Genetic Algorithm (GA) and MARLEDA (a modern Estimation of Distribution Algorithm [56])	Evolutionary Algorithms (Genetic Algorithm)	(Alden, Bryan et al. 2012) [46]
13	Decision Tree (C4.5) + Financial Statements Detection Ontology	Machine Learning (Data Mining) + Knowledge Engineering	(Xiao-Bo, Guang-Chao et al. 2018) [6]
14	Neural Network, Decision Trees, Genetic Algorithms and Bayesian Belief Networks	Machine Learning + Genetic Algorithms	(Albashrawi 2016) [57]
15	Multilayer Feed Forward Neural Network (MLFF), Support Vector Machines (SVM), Genetic Programming (GP), Group Method of Data Handling (GMDH), Logistic Regression (LR), and Probabilistic Neural Network (PNN)	Machine Learning (Data Mining)+ Genetic Algorithm	(Ravisankar, Ravi et al. 2011) [58]

To sum up, the vast majority of studies either do not focus on financial statement fraud or cannot provide sufficient interpretability. For this, an ontology-based framework for financial statements fraud detection was proposed. More importantly, the proposed framework can provide accounts level fraud warning, which is different from most of previous researches. That is, the proposed framework not only can realize the financial statements fraud detection at the company level, but also provides timely risk warning at the accounts level, which is very useful and convenient for regulation, internal control, audit and investment decision.

III. ONTOLOGY-BASED MISSTATEMENT ACCOUNTS DETECTION PROCESS

From the world-beating *Enron Incident* [2] to recent *Luckin Coffee* fraud, financial statements fraud never seems to go far, which confuses auditors, stakeholders and regulators for a long time. Looking back at these cases, we can find that the financial statements fraud is usually rooted from the profits misstatement. Taking one step further, the misstatement of financial statements mainly concentrates on the financial statements items (accounts) closely related to profits. Inspired by this observation, a profit centered financial statements

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3014620, IEEE Access

fraud detection ontology was constructed and used to detect misstatement accounts. In terms of steps (Fig. 1), the ontology-based misstatement accounts detection process includes *ontology construction* and *rule-based fraud detection*, which will be illustrated in detail below.

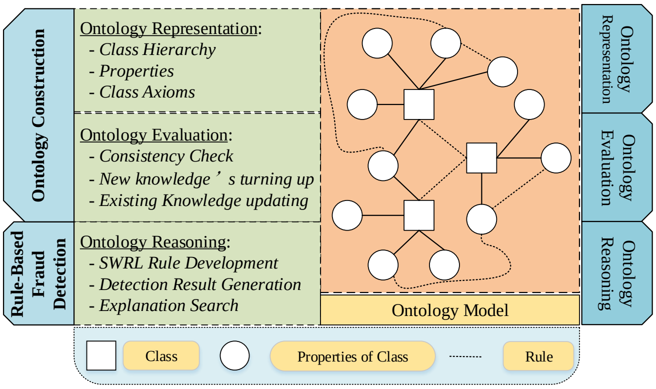

FIGURE 1. Ontology-based misstatements accounts detection process.

A. ONTOLOGY CONSTRUCTION

According to [59], an ontology is a formal representation of concepts and their relations. To be specific, an ontology mainly involves three types knowledge (class axioms, properties, rules) and individuals. Firstly, the class axioms include subclass axioms, equivalent class axioms and disjoint axioms, which is the foundation of class hierarchy. Secondly, the properties can be furtherly divided into data properties and object properties, which can represent quantitative knowledge and qualitative knowledge respectively. Thirdly, the rules, the key knot between ontology and infer engine, are generally used for generating reasoning results. Lastly, the individuals can be regarded as specific instances of the concepts (class). In addition, OWL (*Ontology Web Language*), SWRL (*Semantic Web Rule Language*) [10] are generally used for representing and constructing ontology model.

In light of technical foundation of ontology, knowledge is the foundation and key of ontology construction. Accordingly, the knowledge here mainly involves the domain knowledge that can be used to detect financial statements fraud effectively. Therefore, to better prepare knowledge, it is necessary to make a simple analysis of the financial statements fraud.

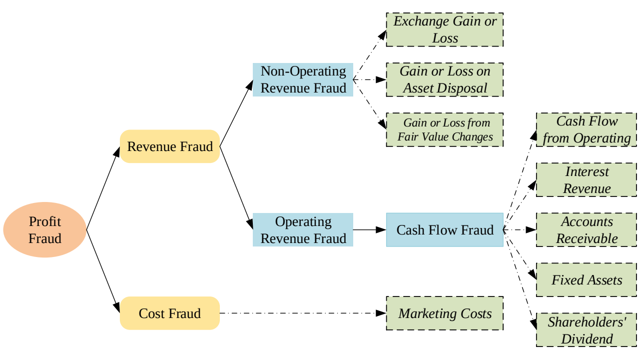

FIGURE 2. Schematic Diagram of Financial Statements Fraud.

TABLE II
INTUITIONS OF FINANCIAL STATEMENTS FRAUD DETECTION

	Intuitions
Intuition 1	*Financial statements fraud is to beautify profits.*
Intuition 2	*Profit fraud usually starts from revenue fraud or cost fraud.*
Intuition 3	*Although cost fraud does not need cash flow as a support, but the space for fraud is relatively limited.*
Intuition 4	*Revenue fraud can be divided into operating revenue fraud and non-operating revenue fraud.*
Intuition 5	*Non-operating revenue fraud is generally adjusted from three aspects: exchange gain or loss, gain or loss on assets disposal and gain or loss from fair value changes.*
Intuition 6	*The realization of operating revenue requires corresponding cash inflow.*
Intuition 7	*In order to cover up the lack of cash inflow, operating revenue fraud usually offsets the cash inflow with accounts receivable, fixed assets and shareholders' dividends.*
Intuition 8	*In addition, occasional projects, such as construction in progress and goodwill, are also the disaster areas of financial statements fraud.*
Intuition 9	*Cash flow is the most difficult to cover up, which will become an important breakthrough in the detection of financial statements fraud.*
Intuition 10	*Behind the abnormal growth of raw financial indicators, there is often misstatement financial statements.*
Intuition 11	*In addition to the raw financial indicators, some financial ratios can also provide clues for the fraud detection of financial statements.*
Intuition 12	*The indicative financial ratios for detect fraud mainly involve non-operating revenue ratio, gross margin, operating margin, cash ratio of net profit, ratio of expenses to sales and return on cash.*

As shown in Fig. 2, the core goal of financial statements fraud is profit fraud, which can be further divided into revenue fraud and cost fraud. That is, financial statements fraud is usually realized by falsely increasing revenue or reducing cost. In fact, the scope of cost reduction is very limited. In this way, most financial statements fraud cases tend to falsely increase revenue. However, the falsely increased revenue should be companied with equivalent cash inflow from operating, which need to be reflected in *Cash Flow Statements*. For this, to cover up the cash flow shortcoming of virtual revenue, past fraud cases usually adopted three kinds of strategies. The first one is to convert the virtual revenue into *Accounts Receivable*, which can be gradually squeezed out through *Loss on Bad Debts*. Secondly, the virtual revenue can be used to *'purchase' Fixed Assets*, which can be disposed of by depreciation at last. Thirdly, the virtual revenue can also be transformed into the

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3014620, IEEE Access

dividend of shareholders, forming a cash flow closed-loop from sales to dividend. Apart from virtual revenue, some companies also increased *Non-Operating Revenue* through disposal of assets, foreign currency business and changes in fair value to make up for the deficit of *Operating Revenue*. Motivated by these observations, 12 intuitions in Table II can be obtained, which constitutes the knowledge base of ontology construction.

For the sake of the professionalism of financial statements fraud detection, Table III shows the implications of 17 raw financial indicators (data properties), which were selected based on the above intuitions and used to build the ontology model.

TABLE III
RAW FINANCIAL INDICATORS OF ONTOLOGY MODEL

Raw Financial Indicators		Implications
1.Balance Sheet Items	*1.1 Cash*	Cash on hand and bank deposits (including cheques and savings account deposits), current cheques and bank drafts.
	1.2 Accounts Receivable	Money to be collected from the purchasing unit for selling goods, products, providing services and other businesses.
	1.3 Inventory	Finished products or commodities held for sale, products in process of production, materials and materials consumed in the process of production or service provision, etc.
	1.4 Fixed Assets	Tangible assets held for the purpose of producing commodities, providing labor services, leasing or business management and with a service life of more than one accounting year.
	1.5 Construction in Progress	Expenditures for new construction, reconstruction and expansion of fixed assets of the company, or unfinished projects such as technical transformation, equipment renewal and major repair projects.
	1.6 Goodwill	The potential economic value that can bring excess profits to the business operation in the future.
2.Income Statement Items	*2.1 Operating Revenue*	In production and operation activities, various revenues from sales of products or provision of services.
	2.2 Operating Profit	Profits realized by the company in all its sales business.
	2.3 Gross Profit	The part of operating revenue after deducting the direct cost of main business.
	2.4 Net Profit	The company's profits are retained after the income tax has been paid in accordance with the provisions in the total profits.
	2.5 Marketing Costs	Various expenses incurred in the process of selling goods and materials and providing services.
	2.6 Interest Revenue	The interest income obtained by the company from the use of funds by others or the occupation of its own funds by others.
	2.7 Exchange Gain or Loss	The difference in the amount of accounting functional currency due to the use of different exchange rates for foreign currency transactions.
	2.8 Gain or Loss on Asset Disposal	Gain or loss from disposal of fixed assets and sale of intangible assets.
	2.9 Gain or Loss from Fair Value Changes	When the fair value measurement mode is adopted subsequently, the difference between the book value of the assets at the end of the period and its fair value.
	2.10 Non-Operating Revenue	Various revenues not directly related to the production and operation of the company.
3.Cash Flow Statement Items	*3.1 Net Cash Flow from Operating*	Cash flow provided by the company after deducting the increase of working capital from the gross cash flow of operating.

It is not enough to effectively detect the misstatement accounts of financial statements only by the 17 raw financial indicators. In view of this, 13 derivative financial ratios (Table IV) were also introduced into the proposed ontology model. With these derivative financial ratios, the articulation across time and statements can be included. In particular, *Difference between Sales and Inventory Growth* was designed by the law of synergies between operating revenue and inventory, which is an innovative indicator in this paper.

Till now, the domain knowledge for ontology construction has been prepared. Taking one step further, with the help of OWL, we can formalize the prepared domain knowledge into the form of class axioms (subclass axioms, equivalent class axioms, disjoint axioms) and properties (data properties and object properties). After that, the next step is the rule-based financial statements fraud detection.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

TABLE IV
DERIVATIVE FINANCIAL INDICATORS FOR FRAUD DETECTION ONTOLOGY

Derivative Indicators	**Financial**	**Implication**	**Formal Definition**
1. Growth Rate	*1.1 Accounts Receivable Growth Rate*	Growth rate of current accounts receivable over the previous period	*(Current Accounts Receivable - Previous Accounts Receivable) / Previous Accounts Receivable*
	1.2 Inventory Growth Rate	Growth rate of current inventory over the previous period	*(Current Inventory - Previous Inventory) / Previous Inventory*
	1.3 Fixed Assets Growth Rate	Growth rate of fixed assets over the previous period	*(Current Fixed Assets - Previous Fixed Assets) / Previous Fixed Assets*
	1.4 Growth Rate of Constructio n in Progress	Growth rate of current construction in progress over the previous period	*(Current Construction in Progress - Previous Construction in Progress) / Previous Construction in Progress*
	1.5 Goodwill Growth Rate	Growth rate of current goodwill over the previous period	*(Current Goodwill - Previous Goodwill) / Previous Goodwill*
	1.6 Growth Rate of Operating Revenue	Growth rate of current operating revenue over the previous period	*(Current Operating Revenue - Previous Operating Revenue) / Previous Operating Revenue*
	1.7 Difference between Sales and Inventory Growth	Difference between operating revenue growth and inventory growth	*Growth Rate of Operating Revenue - Inventory Growth Rate*
2. Financial Ratio	*2.1 Non-Operating Revenue Ratio*	Ratio of non-operating revenue to net profit	*Non-Operating Revenue / Net Profit*
	2.2 Gross Margin	Gross profit as a percentage of operating revenue	*Gross Profit / Operating Revenue*
	2.3 Operating Margin	Ratio of operating profit to operating revenue	*Operating Profit / Operating Revenue*
	2.4 Cash Ratio of Net Profit	The ratio between the net cash flow and the net profit from the current operating activities.	*Net Cash Flow from Operating/Net Profit*
	2.5 Ratio of Expenses to Sales	Ratio of marketing costs to operating revenue	*Marketing Costs / Operating Revenue*
	2.6 Return on Cash	Ratio of interest revenue to cash	*Interest Revenue / Cash*

B. RULE-BASED FRUD DETECTION

To utilize the ontology for detecting financial statements fraud at the account level, some rules should be pre-defined to support ontology reasoning. Inspired by this, to represent the rules for generating the derivative financial ratios and the fraud detection results, SWRL [10] and **Pellet** infer engine were adopted. Specially, **Pellet** infer engine was based on *Tableau* algorithm [60], which can guarantee the correctness and consistency of ontology reasoning. At last, a series of reasoning results for detecting misstatement accounts and risk warning can be generated through ontology reasoning.

IV. METHODOLOGY

According to [59], an complete ontology model should include ontology construction, ontology evaluation and ontology reasoning. To be specific, ontology construction, refers to represent knowledge formally with ontology language, while ontology evaluation is to evaluate the quality of ontology model. Besides, ontology reasoning, is to get reasoning results by infer engine. Based on this, from the perspective of ontology modeling, the ontology-based financial statements fraud detection process can be divided into three steps (Fig. 3): ontology representation, ontology evaluation and ontology reasoning.

FIGURE 3. Steps of Ontology-Based Financial Statements Fraud Detection.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3014620, IEEE Access

A. ONTOLOGY REPRESENTATION

1) CLASS HIERARCHY

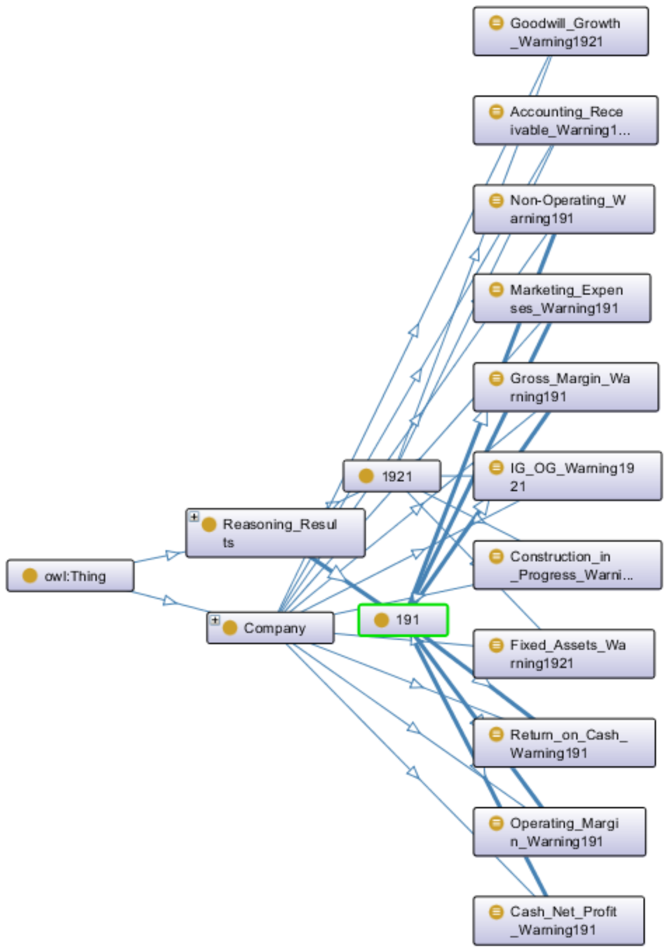

FIGURE 4. The class hierarchy of the constructed ontology.

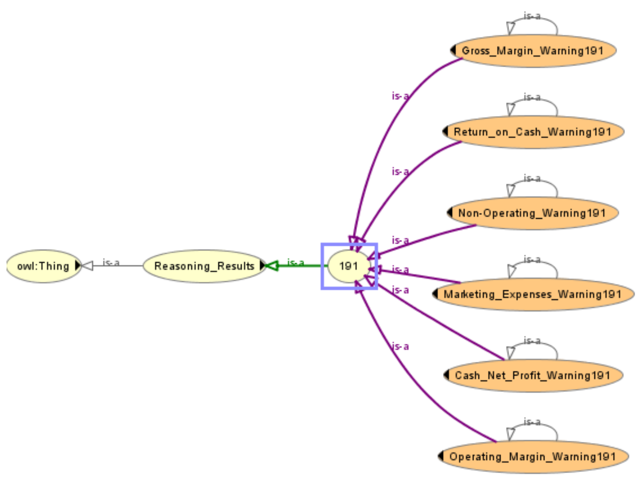

FIGURE 5. The class hierarchy of *Class* "191".

Ontology representation is to formalize the prepared knowledge into the form of class hierarchy (class axioms, properties and rules). In this way, the prepared domain knowledge can be transformed into a structured ontology model, which is the basis of ontology evaluation and ontology reasoning. On the whole, the class hierarchy of the ontology model (Fig. 4) includes the class "*Company*" class and the class "*Reasoning Results*". Taking one step further, the class "*Reasoning Results*" involves five subclasses: "191", "192", "193", "1921", "1932". Taking the class "191" for example, the class hierarchy of the class "191" (financial report fraud warning in the first quarter of 2019) was shown in Fig. 5, where a *'is_a'* corresponds to a subclass axiom or equivalent class axiom.

2) PROPERTIES AND RULES

Furthermore, properties of the ontology model include object properties and data properties. For this framework, data properties were adopted for connecting companies with its raw financial indicators and derivative financial ratios. To be specific, 17 financial indicators (Table III) and 13 derivative financial ratios (Table IV) were introduced to be the data properties of the class. Additionally, rules were represented by SWRL, which were shown in TABLE VIII and used for ontology reasoning.

B. ONTOLOGY EVALUATION

Generally speaking, the goal of ontology evaluation is to verify the consistency and completeness of the ontology model. Accordingly, ontology evaluation here mainly involves consistency check and completeness check.

1) CONSISTENCY CHECK

TABLE V
THE AXIOMS OF CONSISTENCY

Axioms	Formal Definition
Axiom 1	$SubClassOf(C_1, C_2) \wedge SubClassOf(C_2, C_3) \rightarrow SubClassOf(C_1, C_3)$
Axiom 2	$SubClassOf(C_1, C_2) \wedge \dots \wedge SubClassOf(C_i, C_{i+1}) \dots \wedge SubClassOf(C_n, C_1) = False$
Axiom 3	$UnionOfRelation(C_1, C_2) \wedge UnionOfRelation(C_2, C_3) \rightarrow UnionOfRelation(C_1, C_3)$
Axiom 4	$UnionOfRelation(C_1, C_2) \wedge \dots \wedge UnionOfRelation(C_i, C_{i+1}) \dots \wedge UnionOfRelation(C_n, C_1) = False$
Axiom 5	$EquivalentClass(C_1, C_2) \wedge EquivalentClass(C_2, C_3) \rightarrow EquivalentClass(C_1, C_3)$
Axiom 6	$SubClassOf(C_1, C_2) \wedge EquivalentClass(C_2, C_3) \rightarrow SubClassOf(C_1, C_3)$
Axiom 7	$SubClassOf(C_1, C_3) \wedge EquivalentClass(C_1, C_2) \rightarrow SubClassOf(C_2, C_3)$
Axiom 8	$DisjointWith(C_1, C_2) \wedge (SubClassOf(C_3, C_1) \vee EquivalentClass(C_3, C_1)) \wedge (SubClassOf(C_4, C_2) \vee EquivalentClass(C_4, C_2)) \wedge EquivalentClass(C_4, C_2) \rightarrow DisjointWith(C_3, C_4)$

Where *SubClassOf* refers to the subset relationship between concepts; *UnionOfRelation* refers to the whole and part relationship between concepts; *EquivalentClass* refers to the equivalent relationship between concepts; *DisjointWith* refers to the disjoint relationship between concepts.

Ontology consistency involves three meanings: syntax consistency, semantic consistency and domain consistency. Syntax consistency means to satisfy the grammatical rules of

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

OWL. Here, syntax consistency implies that the description of an ontology conforms to the syntactic rules of OWL. Concretely, syntax consistency needs to satisfy *Axiom* 1 to *Axiom* 7 in Table V simultaneously. Semantic consistency, or logical consistency, refers to meet the logical basis of ontology description language, such as *Axiom 8* in Table V. Domain consistency involves specific domain rules. For example, in the financial statements, the total assets are equal to the total liabilities plus the owner's equity. At the technical level, the above three consistency checks can be completed with **Pellet** infer engine and individuals.

2) COMPLETENESS CHECK

Here, the completeness of the ontology model means that, the formal knowledge should be sufficient enough to support financial statements fraud detection. In other words, the domain knowledge introduced should be complete, even a little redundant. Since there is no mature method for completeness check, the completeness check here is mainly based on the consistency of reasoning results and analysis.

C. ONTOLOGY REASONING

In fact, the representation ability of OWL is so limited that it can only express class axioms and properties. As for the rules used for ontology reasoning, it is beyond the representation of OWL. To solve this problem, SWRL was introduced for representing the essential rules. Specifically, in order to generate derivative financial ratios and effective reasoning conclusions, a total of 35 SWRL rules were incorporated into the ontology model.

At last, with the ontology model represented by OWL and SWRL, **Pellet** infer engine can generate misstatement accounts detection results, which can not only detect whether a company has the financial statements fraud detection behavior, but also provide the risk early warning at the account level. With the account level detection results, external auditors, regulators, investors and internal managers can easily further identify disaster areas and causes, which is very convenient and efficient.

V. CASE STUDY

On April 2, 2020, *Luckin Coffee* released a notice admitting that it had made 2.2 billion *yuan* of false transactions, which confuses investors, external auditors, regulators and other stakeholders very much. In response to this problem, *Luckin Coffee* was chosen to be the object of case study. Here, *Company LK* was used as the alternative name of *Luckin Coffee*. Specifically, the process of case study was carried out by ontology construction and rule-based fraud detection.

A. ONTOLOGY CONSTRUCTION

To some extent, the financial statements fraud detection problem, can be defined as the detection of misstatement accounts. For this, the items and articulation in financial statements become the major concern of the proposed framework. Generally speaking, the articulation of financial statements (Fig. 6) is mainly implied between different statements of the same period and the same items of statements of different periods. That is, the articulation of financial statements was mainly reflected in statements dimension and time dimension. In light of this, more than one quarters of financial statements need to be prepared.

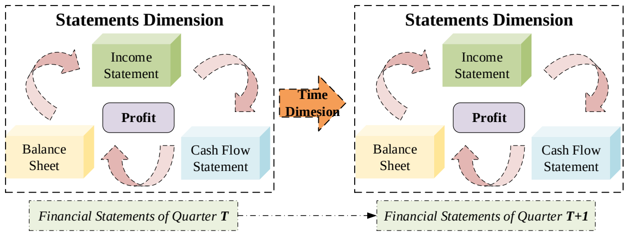

FIGURE 6. The articulation of financial statements.

As we all know, the fraud of *Company LK* was mainly reflected in the financial statements of 2019. Moreover, up to now, the financial statements of *Luckin Coffee* have only been disclosed to the third quarter of 2019. In view of this, the timeframe of the financial statements was set for the first three quarters of 2019. Concretely, 17 raw financial indicators (Table VI) were selected from financial statements. Besides, 13 derivative financial ratios, generated by 17 raw financial indicators (Table IV), were also introduced for detecting financial statements fraud detection. Formally, these financial indicators (ratios) and financial ratios were mainly represented to be data properties by OWL.

TABLE VI
RAW FINANCIAL INDICATORS OF *COMPANY LK*

Raw Financial Indicators		2019 (*million yuan*) Quarter 1	Quarter 2	Quarter 3
1.Balance Sheet Items	*1.1 Cash*	1159	3989	4514
	1.2 Accounts Receivable	7.498	6.95	22.5
	1.3 Inventory	189	231.7	213.2
	1.4 Fixed Assets	966	1064	1238
	1.5 Construction in Progress	--	--	--
	1.6 Goodwill	--	--	--
2.Income Statement Items	*2.1 Operating Revenue*	47.85	909.1	1542
	2.2 Operating Profit	-527.1	-689.7	-590.9
	2.3 Gross Profit	-79.67	71.83	343.2
	2.4 Net Profit	-551.8	-681.3	-531.9
	2.5 Marketing Costs	168.1	390.1	557.7

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

	2.6 Interest Revenue	1.551	14.13	31.85
	2.7 Exchange Gain or Loss	-8.64	5.584	32.8
	2.8 Gain or Loss on Asset Disposal	--	--	--
	2.9 Gain or Loss from Fair Value Changes	-8.322	--	--
	2.10 Non-Operating Revenue	-16.962	5.584	32.8
3.Cash Flow Statement Items	*3.1 Net Cash Flow from Operating*	-627.6	-375.2	-122.8

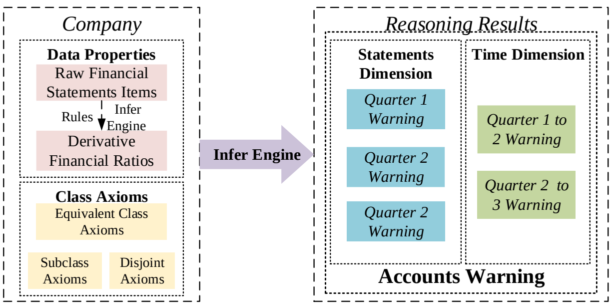

FIGURE 7. Logical Relationships of the Main Classes of the Proposed Ontology.

Taking one step further, combined these financial indicators (ratios) with three types of axioms, the basic architecture of ontology, or class hierarchy (Fig.4) can be formed. As a whole, the logical relationships of the major classes were shown in Fig. 7.

It should be noted that the reasoning is mainly based on equivalence class axioms and rules. Taking the risk warning classed '191' (*Quarter 1 Warning*) and '1921' (*Quarter 1 to 2 Warning*) for example, the relevant equivalent class axioms, used for detecting fraud detection, were shown in Table VII. In particular, '*IG_OG_Warning191*' was creatively introduced for detecting misstatement revenue. It indicates that the difference between operating revenue growth and inventory growth have already exceeded the normal range. To be specific, the value of normal range was from industry practice and empirical study.

TABLE VII
EQUIVALENT CLASS AXIOMS OF CLASS '*191*' AND '*1921*'

	Class	Equivalent Class Axioms
191	Cash_Net_Profit _Warning191	*Company and (Cash_Net_Profit_Ratio191 some xsd:float[< 0.9f])*
	Gross_Margin _Warning191	*Company and (Gross_Margin191 some xsd:float[> 0.5f])*
	Markerting_Cost _Warning191	*Company and (Ratio_of_Marketing_to_Sales191 some xsd:float[> 0.1f])*
	Non-Operating _Warning191	*Company and (Non-Operating_Ratio191 some xsd:float[> 0.1f])*
	Operating_Margin _Warning191	*Company and (Operating_Margin191 some xsd:float[> 0.3f])*
	Return_on_Cash _Warning191	*Company and (Return_on_Cash191 some xsd:float[< 0.015f])*
1921	Accounts_Receivable _Warning1921	*Company and (Accounts_Receivable_Growth1921 some xsd:float[> 0.1f])*
	Construction_in_Prog ress _Warning1921	*Company and (Construction_in_Progress_Growth 1921 some xsd:float[> 0.0f])*
	Fixed_Assets _Warning1921	*Company and (Fixed_Assets_Growth1921 some xsd:float[> 0.1f])*
	Goodwil _Warning1921	*Company and (Goodwill_Growth1921 some xsd:float[> 0.0f])*
	IG_OG _Warning1921	*Company and (IG_OG_Ratio1921 some xsd:float[> 0.1f])*

Apart from class axioms and properties, rules are also the key to ontology reasoning. To represent essential rules, SWRL was introduced. With SWRL, 35 rules were included in the proposed framework. Again, taking class '191' (*Quarter 1 Warning*) and '1921' (*Quarter 1 to 2 Warning*) as examples, 14 representative SWRL rules were listed in Table VIII. Actually, SWRL rules were mainly used to generate derivative financial ratios. Taking SWRL rule 'R1' for example, *Company (?A) ^ Interest_Revenue191 (?A, ?M) ^ Cash191 (?A, ?N) ^ swrlb: divide (?O, ?M, ?N) -> Return_on_Cash191 (?A, ?O)* implies that, in the first quarter of 2019, the *Return on Cash* of *Company LK* is equal to the corresponding interest revenue divided by the amount of cash.

After constructing the ontology model, the class axioms, properties and rules were all ready for ontology evaluation and ontology reasoning. In fact, consistency checking is done synchronously with reasoning. Besides, the proof of completeness has been reflected in the articulation between different items of financial statements and different reporting periods. In light of this, ontology reasoning was left alone.

TABLE VIII
SWRL RULES OF CLASS '*191*' AND '*1921*'

		SWRL Rules
191	R1	*Company(?A) ^ Interest_Revenue191(?A, ?M) ^ Cash191(?A, ?N) ^ swrlb:divide(?O, ?M, ?N) -> Return_on_Cash191(?A, ?O)*

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3014620, IEEE Access

	R2	*Company(?A) ^ Net_Cash_Flow_from_Operating191(?A, ?M) ^ Net_Profit191(?A, ?N) ^ swrlb:divide(?O, ?M, ?N) -> Cash_Net_Profit_Ratio191(?A, ?O)*
	R3	*Company(?A) ^ Gross_Profit191(?A, ?M) ^ Operating_Revenue191(?A, ?N) ^ swrlb:divide(?O, ?M, ?N) -> Gross_Margin191(?A, ?O)*
	R4	*Company(?A) ^ Operating_Profit191(?A, ?M) ^ Operating_Revenue191(?A, ?N) ^ swrlb:divide(?O, ?M, ?N) -> Operating_Margin191(?A, ?O)*
	R5	*Company(?A) ^ Marketing_Costs191(?A, ?M) ^ Operating_Revenue191(?A, ?N) ^ swrlb:divide(?O, ?M, ?N) -> Ratio_of_Marketing_to_Sales191(?A, ?O)*
	R6	*Company(?A) ^ autogen0:Loss191(?A, ?M) ^ autogen1:Loss191(?A, ?N) ^ Changes_in_Fair_Value191(?A, ?O) ^ swrlb:add(?P, ?M, ?N, ?O) -> Non-Operating_Revenue191(?A, ?P)*
	R7	*Company(?A) ^ Non-Operating_Revenue191(?A, ?M) ^ Net_Profit191(?A, ?N) ^ swrlb:divide(?O, ?M, ?N) -> Non-Operating_Ratio191(?A, ?O)*
1921	S1	*Company(?A) ^ Accounts_Receivable192(?A, ?M) ^ Accounts_Receivable191(?A, ?N) ^ swrlb:subtract(?O, ?M, ?N) ^ swrlb:divide(?P, ?O, ?N) -> Accounts_Receivable_Growth1921(?A, ?P)*
	S2	*Company(?A) ^ Fixed_Assets192(?A, ?M) ^ Fixed_Assets191(?A, ?N) ^ swrlb:subtract(?O, ?M, ?N) ^ swrlb:divide(?P, ?O, ?N) -> Fixed_Assets_Growth1921(?A, ?P)*
	S3	*Company(?A) ^ Inventory192(?A, ?M) ^ Inventory191(?A, ?N) ^ swrlb:subtract(?O, ?M, ?N) ^ swrlb:divide(?P, ?O, ?N) -> Inventory_Growth1921(?A, ?P)*
	S4	*Company(?A) ^ Operating_Revenue192(?A, ?M) ^ Operating_Revenue191(?A, ?N) ^ swrlb:subtract(?O, ?M, ?N) ^ swrlb:divide(?P, ?O, ?N) -> Operating_Revenue_Growth1921(?A, ?P)*
	S5	*Company(?A) ^ Operating_Revenue_Growth1921(?A, ?M) ^ Inventory_Growth1921(?A, ?N) ^ swrlb:subtract(?O, ?M, ?N) -> IG_OG_Ratio1921(?A, ?O)*
	S6	*Company(?A) ^ Goodwill_192(?A, ?M) ^ Goodwill_191(?A, ?N) ^ swrlb:subtract(?O, ?M, ?N) -> Goodwill_Growth1921(?A, ?O)*
	S7	*Company(?A) ^ Construction_in_Progress192(?A, ?M) ^ Construction_in_Progress191(?A, ?N) ^ swrlb:subtract(?O, ?M, ?N) -> Construction_in_Progress_Growth1921(?A, ?O)*

B. RULE-BASED FRAUD DETECTION

With the formal ontology and **Pellet** infer engine, the major reasoning results of *Company LK* (Fig. 8) can be generated. After reasoning, 13 pieces of accounts warning across three quarters were generated, which is basically consistent with the misstatement disclosure of *Company LK*. This shows the reliability and validity of the proposed framework. Notably, to the best of our knowledge, this is also the first framework for detecting financial statements fraud at the account level.

More importantly, the reasoning results can not only focus on specific subjects, but also provide sufficient explanation for each reasoning result. This is what machine learning, especially deep learning, cannot do. Taking *IG_OG_Warning1921* as an example, the **Pellet** infer engine can provide a total of 45 reasoning explanations. Given the limited space, Fig. 9 shows only the first three reasoning explanations. Among these explanations, the first one runs through the whole process from the raw financial indicators, to the derivative financial ratios and equivalent class axioms, and finally to the reasoning results.

FIGURE 8. Reasoning Results of *Company LK*.

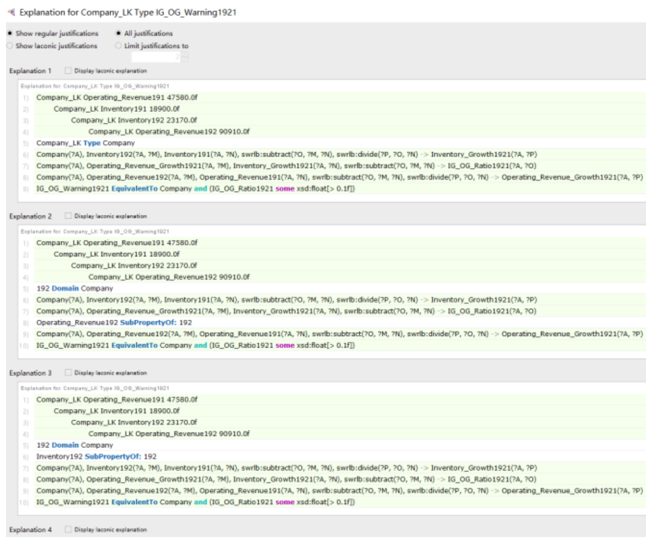

FIGURE 9. Three Reasoning Explanations for *IG_OG_Warning1921*.

Based on the case study above, the effectiveness of the proposed framework has been fully verified. What's more, it has strong interpretability for specific accounts warning. Notably, to the best of our knowledge, this is also the first framework for detecting financial statements fraud at the account level.

V. CONCLUSION

Aiming at achieve financial statements fraud detection at the account level, an ontology-based framework was proposed. The biggest difference between the framework and the previous research is that it not only realizes the financial statements fraud detection at accounts level, but also provides enough explanations of accounts warning, which is very useful for regulation, internal control, audit and risk early warning.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3014620, IEEE Access

Moreover, the case study has fully illustrated the effectiveness and practicability of the proposed framework. In particular, our framework mainly includes the following 5 contributions.

1) ACCOUNTS LEVEL FRAUD DETECTION FOR THE FIRST TIME

In the past, the intelligent financial statements fraud detection can only be accurate to the company level. That is, they can only judge whether a company has financial statements fraud, but cannot give the specific fraudulent accounts. To the best of our knowledge, the proposed framework is almost the first one accurate to the accounts level. This is of great significance to internal control personnel, regulators, auditors, investors and other stakeholders.

2) INTERPRETABLE AND TIMELY WARNING

Compared with most machine learning methods, ontology model has stronger interpretability and timeliness. For one thing, the knowledge of machine learning mainly comes from big data, and its theoretical explanation is not yet fully mature. The natural interpretability of ontology model makes up for this, and does not need millions of data. For another thing, machine learning models usually have longer running time than ontology reasoning. Therefore, ontology model is more suitable for fraud detection of financial statements. After all, the earlier financial statements fraud is identified, the smaller the loss.

3) COMBINATION OF TIME DIMENSION AND STATEMENTS DIMENSION

The clues of financial statements fraud are usually hidden in different period statements and different statements items. Based on this observation, the proposed framework creatively realizes the organic combination of time dimension clues and statements dimension clues through derivative financial ratios. What's more, the case study shows its effectiveness for financial statements fraud detection.

4) COLLABORATIVE FILTERING OF INVENTORY AND OPERATING REVENUE

In view of the persistent problem of operating revenue fraud, apart from the common financial ratios, *Difference in Sales and Inventory Growth* was also creatively designed and introduced for fraud detection. The rationality of these financial ratios lies in that, for the retail companies selling physical goods, the increase of operating revenue should be accompanied by the almost synchronous increase of inventory. Therefore, when the operating revenue of the company increases rapidly, but the inventory does not increase synchronously, the company is very likely to have the fraud behavior of falsely increasing the operating revenue.

5) FLEXIBLE AND ADJUSTABLE MECHANISM BASED ON THRESHOLDS

The actual situation of different industries and companies of different scales is often different, so the normal range of their financial indicators (ratios) is usually variable. To solve this problem, threshold was introduced into the proposed framework. In this way, even in the face of different companies, personalized ontology model can be customized through threshold fine-tuning.

Despite of this, there are still many limitations in our research. For example, the proposed framework needs to be further verified, more domain knowledge for financial statements fraud detection still lack, and the reasoning results needs to be further integrated. In light of this, there are still some problems need to be solved in the future work.

REFERENCES

[1] N. Mohamed, "Financial Statements Fraud Control: Exploring Internal Control Strategies in Two Malaysian Public Interest Entities'," 2010.

[2] Saarni and Jenna, "Financial Fraud - Importance of an Internal Control System," *Haaga Helia Ammattikorkeakoulu,* vol. 39, no. 3, pp. 355-366, 2012.

[3] G. Dickey, S. Blanke, and L. Seaton, "Machine Learning in Auditing," *CPA Journal,* Article vol. 89, no. 6, pp. 16-21, 2019.

[4] M. Koenig, "Knowledge Management in Theory and Practice (2nd ed.)," *JASIST,* vol. 62, p. 2083, 10/01 2011.

[5] D. Zhou, L. Zheng, Y. Zhu, J. Li, and J. He, "Domain Adaptive Multi-Modality Neural Attention Network for Financial Forecasting," presented at the Proceedings of The Web Conference 2020, Taipei, Taiwan, 2020.

[6] T. Xiao-Bo, L. Guang-Chao, Y. Jing, and W. Wei, "Knowledge-based Financial Statement Fraud Detection System: Based on an Ontology and a Decision Tree," *Knowledge Organization,* Article vol. 45, no. 3, pp. 205-219, 2018.

[7] H. Dan, "Researches of Detection of Fraudulent Financial Statements based on Data Mining," *Revista de la Facultad de Ingenieria,* Article vol. 32, no. 5, pp. 463-469, 2017.

[8] S. Huang, C.-C. Lin, A.-A. Chiu, and D. Yen, "Fraud detection using fraud triangle risk factors," *Information Systems Frontiers,* Article vol. 19, no. 6, pp. 1343-1356, 12// 2017.

[9] J. Nigrini Mark, "The patterns of the numbers used in occupational fraud schemes," *Managerial Auditing Journal,* vol. 34, no. 5, pp. 606-626, 2019.

[10] Horrocks *et al.*, "SWRL: A Semantic Web rule language combining OWL and RuleML," *W3C Subm,* vol. 21, 01/01 2004.

[11] A. Martin, M. Manjula, and D. V. P. Venkatesan, "A Business Intelligence Model to Predict Bankruptcy using Financial Domain Ontology with Association Rule Mining Algorithm," ed, 2011.

[12] J. Organ and L. Stapleton, "The Control of Human Factors in Catastrophic Financial Systems Risk using Ontologies," *IFAC PapersOnLine,* Article vol. 50, no. 1, pp. 6367-6372, 2017.

[13] A. Pătraşcu, "Ontology based Approach for an Insurance Company Activity Modelling," *Ovidius University Annals, Series Economic Sciences,* vol. 16, no. 1, p. 81, 2016.

[14] Y. Weir, A. Sujanani, P. Ray, N. Paramesh, D. Lee, and R. Bhar, "DESIGN AND DEVELOPMENT OF FINANCIAL APPLICATIONS USING ONTOLOGY-BASED MULTI-AGENT SYSTEMS," *Computing & Informatics,* vol. 28, no. 5, p. 635, 2009.

[15] W. Ying, P. Ray, and L. Lewis, "A Methodology for Creating Ontology-Based Multi-agent Systems with an Experiment in Financial Application Development," in *2013 46th Hawaii International Conference on System Sciences*, 2013, pp. 3397-3406.

[16] M. d. P. Salas-Zárate, R. Valencia-García, A. Ruiz-Martínez, and R. Colomo-Palacios, "Feature-based opinion mining in financial news: An ontology-driven approach," *Journal of Information Science,* Article vol. 43, no. 4, pp. 458-479, 2017.

[17] P. I. Dissanayake, T. K. Colicchio, and J. J. Cimino, "Using clinical reasoning ontologies to make smarter clinical decision support systems: a systematic review and data synthesis," *Journal of the American Medical Informatics Association,* Article vol. 27, no. 1, pp. 159-174, 2020.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3014620, IEEE Access

[18] S. El-Sappagh, J. M. Alonso, F. Ali, A. Ali, J. Jang, and K. Kwak, "An Ontology-Based Interpretable Fuzzy Decision Support System for Diabetes Diagnosis," *IEEE Access, Access, IEEE,* Periodical vol. 6, pp. 37371-37394, 2018.

[19] Y. Shen *et al.*, "An ontology-driven clinical decision support system (IDDAP) for infectious disease diagnosis and antibiotic prescription," *Artificial Intelligence In Medicine,* Article vol. 86, pp. 20-32, 2018.

[20] Y. Shen, J.-A. Armelle, and J. Colloc, "A multi-agent ontologies-based clinical decision support system," ed, 2020.

[21] Y. Shen, J. Colloc, A. Jacquet-Andrieu, Z. Guo, and Y. Liu, "Constructing Ontology-Based Cancer Treatment Decision Support System with Case-Based Reasoning," ed, 2018.

[22] J.-B. Lamy, K. Sedki, and R. Tsopra, "Explainable decision support through the learning and visualization of preferences from a formal ontology of antibiotic treatments," *Journal of Biomedical Informatics,* Article vol. 104, 2020.

[23] S. Singh, S. Ghosh, J. Jayaram, and M. K. Tiwari, "Enhancing supply chain resilience using ontology-based decision support system," *International Journal of Computer Integrated Manufacturing,* Article vol. 32, no. 7, pp. 642-657, 2019.

[24] X. Wang, T. N. Wong, and Z.-P. Fan, "Ontology-based supply chain decision support for steel manufacturers in China," *Expert Systems with Applications,* vol. 40, no. 18, pp. 7519-7533, 2013/12/15/ 2013.

[25] M. M. Mabkhot, S. K. Amri, S. Darmoul, A. M. Al-Samhan, and S. Elkosantini, "An ontology-based multi-criteria decision support system to reconfigure manufacturing systems," *IISE Transactions,* Article vol. 52, no. 1, pp. 18-42, 2020.

[26] S. Zillner, A. Ebel, and M. Schneider, "Towards intelligent manufacturing, semantic modelling for the steel industry," *IFAC-PapersOnLine,* vol. 49, no. 20, pp. 220-225, 2016.

[27] Q. Bao, J. Wang, and J. Cheng, "Research on ontology modeling of steel manufacturing process based on big data analysis," in *7th International Conference on Mechatronics and Manufacturing, ICMM 2016, January 13, 2016 - January 14, 2016*, Singapore, Singapore, 2016, vol. 45: EDP Sciences.

[28] M. M. Mabkhot, A. M. Al-Samhan, and L. Hidri, "An Ontology-Enabled Case-Based Reasoning Decision Support System for Manufacturing Process Selection," *Advances in Materials Science & Engineering,* Article pp. 1-18, 2019.

[29] F. Z. Xing, E. Cambria, and R. E. Welsch, "Natural language based financial forecasting: a survey," *Artificial Intelligence Review,* Article vol. 50, no. 1, pp. 49-73, 06// 2018.

[30] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning : A systematic literature review: 2005–2019," *Applied Soft Computing,* vol. 90, p. 106181, 2020/05/01/ 2020.

[31] O. Bustos and A. Pomares-Quimbaya, "Stock market movement forecast: A Systematic review," *Expert Systems with Applications,* vol. 156, p. 113464, 2020/10/15/ 2020.

[32] J. Cao, Z. Li, and J. Li, "Financial time series forecasting model based on CEEMDAN and LSTM," *Physica A: Statistical Mechanics and its Applications,* vol. 519, pp. 127-139, 2019/04/01/ 2019.

[33] B. S. Kumar and V. Ravi, "A survey of the applications of text mining in financial domain," *Knowledge-Based Systems,* vol. 114, pp. 128-147, 2016/12/15/ 2016.

[34] H. Naderi Semiromi, S. Lessmann, and W. Peters, "News will tell: Forecasting foreign exchange rates based on news story events in the economy calendar," *The North American Journal of Economics and Finance,* vol. 52, p. 101181, 2020/04/01/ 2020.

[35] R. Chiong, Z. Fan, Z. Hu, M. T. P. Adam, B. Lutz, and D. Neumann, "A sentiment analysis-based machine learning approach for financial market prediction via news disclosures," presented at the Proceedings of the Genetic and Evolutionary Computation Conference Companion, Kyoto, Japan, 2018.

[36] J. B. Heaton, N. G. Polson, and J. H. Witte, "Deep Learning in Finance," ed, 2016.

[37] L. Zhang, C. Aggarwal, and G.-J. Qi, "Stock Price Prediction via Discovering Multi-Frequency Trading Patterns," presented at the Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 2017.

[38] O. B. Sezer and A. M. Ozbayoglu, "Algorithmic Financial Trading with Deep Convolutional Neural Networks: Time Series to Image Conversion Approach," *Applied Soft Computing,* p. S1568494618302151, 2018.

[39] F. Kamalov, "Forecasting significant stock price changes using neural networks," ed, 2019.

[40] Z. Zhou, M. Gao, Q. Liu, and H. Xiao, "Forecasting stock price movements with multiple data sources: Evidence from stock market in China," *Physica A: Statistical Mechanics and its Applications,* vol. 542, p. 123389, 2020/03/15/ 2020.

[41] N. S. Gill and R. Gupta, "Analysis of Data Mining Techniques for Detection of Financial Statement Fraud," *IUP Journal of Systems Management,* Article vol. 10, no. 1, pp. 7-15, 2012.

[42] G. L. Gray and R. S. Debreceny, "A taxonomy to guide research on the application of data mining to fraud detection in financial statement audits," *International Journal of Accounting Information Systems,* vol. 15, no. 4, pp. 357-380, 2014/12/01/ 2014.

[43] J. West and M. Bhattacharya, "Mining Financial Statement Fraud: An Analysis of Some Experimental Issues," ed, 2015.

[44] J. Perols, "Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms," *Auditing: A Journal of Practice & Theory,* Article vol. 30, no. 2, pp. 19-50, 2011.

[45] B. Hoogs, T. Kiehl, C. Lacomb, and D. Senturk, "A genetic algorithm approach to detecting temporal patterns indicative of financial statement fraud," *Intelligent Systems in Accounting, Finance & Management,* Article vol. 15, no. 1/2, pp. 41-56, 2007.

[46] M. E. Alden, D. M. Bryan, B. J. Lessley, and A. Tripathy, "Detection of Financial Statement Fraud Using Evolutionary Algorithms," *Journal of Emerging Technologies in Accounting,* Article vol. 9, pp. 71-94, 2012.

[47] D. Wang *et al.*, "A Semi-supervised Graph Attentive Network for Financial Fraud Detection," ed, 2020.

[48] A. Abbasi, C. Albrecht, A. Vance, and J. Hansen, "METAFRAUD: A META-LEARNING FRAMEWORK FOR DETECTING FINANCIAL FRAUD," *MIS Quarterly,* Article vol. 36, no. 4, pp. 1293-A12, 12// 2012.

[49] L. Tang, P. Peng, and C. Luo, "FINANCIAL STATEMENT FRAUD DETECTION THROUGH MULTIPLE INSTANCE LEARNING," *Обнаружение фальсификаций финансовой отчетности через многовариантное обучение,* Article no. 3, pp. 146-155, 06// 2016.

[50] M. Omidi, Q. Min, V. Moradinaftchali, and M. Piri, "The Efficacy of Predictive Methods in Financial Statement Fraud," *Discrete Dynamics in Nature & Society,* Article pp. 1-12, 2019.

[51] J. L. Perols, "Detecting financial statement fraud: Three essays on fraud predictors, multi-classifier combination and fraud detection using data mining," ed, 2008.

[52] C. Seiffert, T. M. Khoshgoftaar, J. V. Hulse, and A. Napolitano, "RUSBoost: A Hybrid Approach to Alleviating Class Imbalance," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans,* vol. 40, no. 1, pp. 185-197, 2010.

[53] P. Hajek and R. Henriques, "Mining corporate annual reports for intelligent detection of financial statement fraud - A comparative study of machine learning methods," vol. 128, ed, 2017, pp. 139-152.

[54] Y.-J. Chen, W.-C. Liou, Y.-M. Chen, and J.-H. Wu, "Fraud detection for financial statements of business groups," *International Journal of Accounting Information Systems,* vol. 32, pp. 1-23, 2019/03/01/ 2019.

[55] S. Xinping and G. Yan, "Detecting Financial Statement Fraud: a Comparative Study Using Data Mining Methods," *International Review on Computers & Software,* Article vol. 7, no. 4, pp. 1778-1783, 2012.

[56] M. Alden and R. Miikkulainen, "MARLEDA: Effective distribution estimation through Markov random fields," *Theoretical Computer Science,* vol. 633, pp. 4-18, 2016/06/20/ 2016.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3014620, IEEE Access

[57] M. Albashrawi, "Detecting Financial Fraud Using Data Mining Techniques: A Decade Review from 2004 to 2015," *Journal of Data Science,* Article vol. 14, no. 3, pp. 553-569, 2016.

[58] P. Ravisankar, V. Ravi, G. R. Rao, and I. Bose, "Detection of financial statement fraud and feature selection using data mining techniques," vol. 50, ed, 2011, pp. 491-500.

[59] V. DevedŽIĆ, "Understanding Ontological Engineering," *Communications of the ACM*, vol. 45, no. 4, p. 136, 04// 2002.

[60] F. Baader and U. Sattler, "Tableau Algorithms for Description Logics," Berlin, Heidelberg, 2000, pp. 1-18: Springer Berlin Heidelberg.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.