Techscience IT Magazine

DEPARTMENT OF IT

July 18- December 18

Editors

Prof. Saritha L R

Prof. Savita Lohiya

Where Is Current Research on Blockchain Technology?

Currency transactions between persons or companies are often centralized and controlled by a third party organization. Making a digital payment or currency transfer requires a bank or credit card provider as a middle man to complete the transaction. In addition, a transaction causes a fee from a bank or a credit card company. The same process applies also in several other domains, such as games, music, software etc. The transaction system is typically centralized, and all data and information are controlled and managed by a third party organization ,rather than the two principal entities involved in the transaction. Blockchain technology has been developed to solve this issue. The goal of Blockchain technology is to create a decentralized environment where no third party is in control of the transactions and data. Blockchain is a distributed database solution that maintains a continuously growing list of data records that are confirmed by the nodes participating in it. The data is recorded in a public ledger, including information of every transaction ever completed. Blockchain is a decentralized solution which does not require any third party organization in the middle. The information about every transaction ever completed in Blockchain is shared and available to all nodes. This attribute makes the system more transparent than centralized transactions involving a third party.

In addition, the nodes in Blockchain are all anonymous, which makes it more secure for other nodes to confirm the transactions. Bitcoin was the first application that introduced Blockchain technology. Bitcoin created a decentralized environment for crypto currency, where the participants can buy and exchange goods with digital money. However, even though Blockchain seems to be a suitable solution for conducting transactions by using crypto currencies, it has still some technical challenges and limitations that need to be studied and addressed. High integrity of transactions and security, as well as privacy of nodes are needed to prevent attacks and attempts to disturb transactions in Blockchain. In addition, confirming transactions in the Blockchain requires a computational power. It is important to identify what topics have been already studied and addressed in Block-chain and what are currently the biggest challenges and limitations that need further studies. To address these questions, we decided to use a systematic mapping study process to identify relevant papers related to Blockchain. In the systematic mapping study, we applied a well-designed research protocol to search for material in scientific databases.

The produced map of current research on Blockchain will help other researchers and practitioners in identifying possible research areas and questions for future research. Although crypto currencies are also a business and management topic, we decided to narrow down the research topic to the technical perspective of Blockchain. Our objective was to find and map all papers with technical viewpoints on Blockchain. We were interested in finding Blockchain research topics related to various technical areas, such as security, performance, data integrity, privacy, and scalability. While blockchain isn't going to replace traditional corporate relational databases, it does open new doors for the movement and storage of transactional data inside and outside of global enterprises.

Blockchain adoption is expected be steady, as the changes it brings gain momentum, according Karim Lakhani, a principal investigator of the Crowd Innovation Lab and NASA Tournament Lab at the Harvard Institute for Quantitative Social Science. "Conception ally, this is TCP/IP applied to the world of business and transactions," Lakhani said. "In the '70s and '80s, TCP/IP was not imaginable to be as robust and scalable as it was. Now, we know that TCP/IP allows us all this modern functionality that we take for granted on the web.

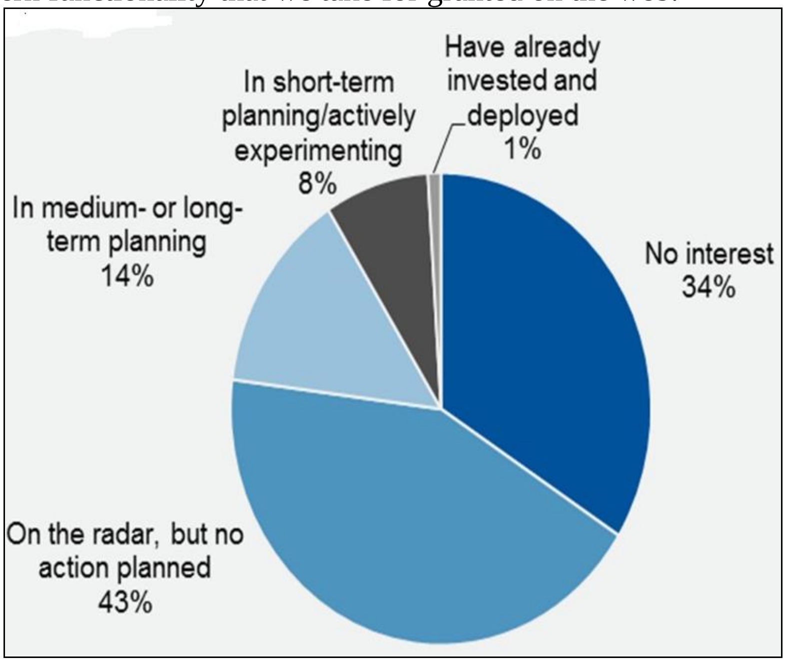

Figure 1 : Blockchain Plan

They are engaging a result of their capacity to recognize zero-day assaults. Another preferred standpoint is that the profiles of typical movement are tweaked for each system, application, or system, along these lines making it troublesome for assailants to know which exercises they can complete undetected. Furthermore, the information on which abnormality based systems caution can be utilized to characterize the marks for abuse finders. The fundamental hindrance of anomaly based methods is the potential for high false alert rates on the grounds that already concealed system practices might be ordered as oddities. This centers essentially around digital interruption detection as it applies to wired systems. With a wired system, a foe must go through a few layers of safeguard at firewalls and working systems, or increase physical access to the system. Nonetheless, a remote system can be focused at any hub, so it is normally more defenseless against pernicious assaults than a wired system.

A Few Useful Things to Know About Machine Learning

Machine learning systems automatically learn programs from data. This is often a very attractive alternative to manually constructing them, and in the last decade the use of machine learning has spread rapidly throughout computer science and beyond. Machine learning is used in Web search, spam filters, recommender systems, ad placement, credit scoring, fraud detection, stock trading, drug design, and many other applications. A recent report from the McKinsey Global Institute asserts that machine learning (a.k.a. data mining or predictive analytics) will be the driver of the next big wave of innovation. Several fine textbooks are available to interested practitioners and researchers (for example, Mitchell16 and Witten et al.24). However, much of the "folk knowledge" that is needed to successfully develop machine learning applications is not readily available in them. As a result, many machine learning projects take much longer than necessary or wind up producing less-than-ideal results. Yet much of this folk knowledge is fairly easy to communicate.

Many different types of machine learning exist, but for illustration purposes I will focus on the most mature and widely used one: classification. Nevertheless, the issues I will discuss apply across all of machine learning. A classifier is a system that inputs (typically) a vector of discrete and/or continuous feature values and outputs a single discrete value, the class. For example, a spam filter classifies email messages into "spam" or "not spam," and its input may be a Boolean vector x = (x1,…,xj,…,xd), where xj = 1 if the j th word in the dictionary appears in the email and xj = 0 otherwise. A learner inputs a training set of examples (xi, yi), where xi = (xi,1 , . . . , xi,d) is an observed input and yi is the corresponding output, and outputs a classifier. The test of the learner is whether this classifier produces the correct output yt for future examples xt (for example, whether the spam filter correctly classifies previously unseen email messages as spam or not spam). Learning = Representation + Evaluation + Optimization Suppose you have an application that you think machine learning might be good for. The first problem facing you is the bewildering variety of learning algorithms available. Which one to use? There are literally thousands available, and hundreds more are published each year. The key to not getting lost in this huge space is to realize that it consists of combinations of just three components. The components are: ¸ Representation. A classifier must be represented in some formal language that the computer can handle. Conversely, choosing a representation for a learner is tantamount to choosing the set of classifiers that it can possibly learn. This set is called the hypothesis space of the learner. If a classifier is not in the hypothesis space, it cannot be learned. A related question, that I address later, is how to represent the input, in other words, what features to use. ¸ Evaluation. An evaluation function (also called objective function or scoring function) is needed to distinguish good classifiers from bad ones. The evaluation function used internally by the algorithm may differ from the external one that we want the classifier to optimize, for ease of optimization .

Optimization. Finally, we need a method to search among the classifiers in the language for the highest-scoring one. The choice of optimization technique is key to the efficiency of the learner, and also helps determine the classifier produced if the evaluation function has more than one optimum. It is common for new learners to start out using off-the-shelf optimizers, which are later replaced by custom-designed ones. The accompanying table shows common examples of each of these three components. For example, k nearest neighbor classifies a test example by finding the k most similar training examples and predicting the majority class among them. Hyper I plane-based methods form a linear mage by agsandrew/Shutterstock.com articles combination of the features per class and predict the class with the highest-valued combination. Decision trees test one feature at each internal node, with one branch for each feature value, and have class predictions at the leaves. Algorithm 1 (above) shows a bare-bones decision tree learner for Boolean domains, using information gain and greedy search. Info Gain(xj, y) is the mutual information between feature xj and the class y. Make Node(x,c0,c1) returns a node that tests feature x and has c0 as the child for x = 0 and c1 as the child for x = 1. Of course, not all combinations of one component from each column of the table make equal sense. For example, discrete representations naturally go with combinatorial optimization, and continuous ones with continuous optimization. Nevertheless, many learners have both discrete and continuous components, and in fact the day may not be far when every single possible combination has appeared in some learner! Most textbooks are organized by representation, and it is easy to overlook the fact that the other components are equally important. There is no simple recipe for choosing each component, but I will touch on some of the key issues here. As we will see, some choices in a machine learning project may be even more important than the choice of learner.