PennState College of Engineering NEWS

< Back

The team hopes to prevent future data attacks for manufacturing companies by providing this innovative methodology. IMAGE: ISTOCK/@MF3D

New method enables automated protections for sensitive data

10/5/2020

By Miranda Buckheit

UNIVERSITY PARK, Pa. — Just as people need to protect their sensitive data, such as social security numbers, manufacturing companies need to protect their sensitive corporate data. There are currently fewer protections for proprietary manufacturing information, making it a ripe environment for corporate data theft of such things as design models.

A particular approach known as differential privacy may be able to better preserve a manufacturer's business, sensitive design details and overall company reputation, a team of Penn State researchers and graduate students report in the Journal of Smart and Sustainable Manufacturing of the **American Society for Testing and Materials**.

"Cyberattacks are increasingly seen in manufacturing," said Hui Yang, professor of industrial engineering. "This brings unexpected disruptions to routine operations and causes the loss of billions of dollars. For example, adversaries often attempt to infer samples included in the training dataset used to create an analytical model or use the released model to infer sensitivity of a target when other background information about this target is available. As manufacturing systems are the backbone of a nation's critical infrastructure for economic growth, there is an urgent need to protect privacy information of manufacturing enterprises and minimize the risk of model inversion attacks."

Companies often data mine large datasets to understand patterns that could increase profits, lower costs, reduce risks and more. Data mining can inadvertently expose private data, posing significant security threats to manufacturers because confidential data such as

customers' identities, production specifications and confidential business information may be compromised.

Differential privacy is an emerging approach to safeguard data from any attempt that may reveal any sensitive data within a system. Differential privacy can fix this problem by creating a scheme that forces the system to create "noise" around the data that needs most protection and by optimizing the privacy parameters for these different kinds of data.

"The idea of preserving privacy was already present, but it gets much more attention now," said Soundar Kumara, the Allen E. Pearce and Allen M. Pearce Professor of Industrial Engineering. "Differential privacy methods are able to put measurements on how much privacy is needed in various scenarios, which is greatly useful for companies. Some information simply isn't as sensitive, like a pet's name versus credit card information. There are applications aimed at differential privacy for smart manufacturing and data mining, and our proposed methodology shows great potential to be applicable for data-enabled, smart and sustainable manufacturing."

The researchers carefully calibrated a model with noise for specific, more sensitive kinds of raw data. The curated, regulated noise contains numerical values that sit among the real information to create distractions, or randomness, within the system to blur what an attacker may see.

The group used test data to evaluate and validate the proposed privacy-preserving data mining framework. They specifically focused on power consumption modeling in computer numerical control (CNC) turning processes.

According to the team, the CNC turning is a precise and intricate manufacturing process in which a rotating workpiece is held in place while a cutter shapes the material. This kind of information can be critical for a manufacturing company, because it may be for their specific product in a competitive market.

"A simple example is a hospital with 500 patients where medical treatments are guided by data mining models trained with their genotype and demographic background," said Qianyu Hu, an industrial engineering doctoral candidate. "If someone outside of the system wants to know specific attributes on patients, for example, their genetic markers, they will attack the model. With normal data, unprotected by noise, an attacker with some background information is able to gain knowledge of the genomic attributes of patients. This knowledge can be adversely used against them in various ways. In this example, adding noise to the data mining process, based on our model, can lower the risk of privacy leakage."

The team noted that in their future research, they plan to continue testing the proposed data mining framework to a network of collaborative manufacturers.

Ruimin Chen, industrial engineering doctoral candidate, also contributed to this work.

The authors received support from Yang's National Science Foundation CAREER grant (CMMI-1617148) for this research.

Share this story:

MEDIA CONTACT:

Megan Lakatos

mkl5024@psu.edu

RELATED STORIES:

- Algorithm aims to alert consumers before they use illicit online pharmacies
- New method analyzes images to improve health care and manufacturing