JuicyScore

Device ID true power in the digital world

Amplifying the potential of Fintech

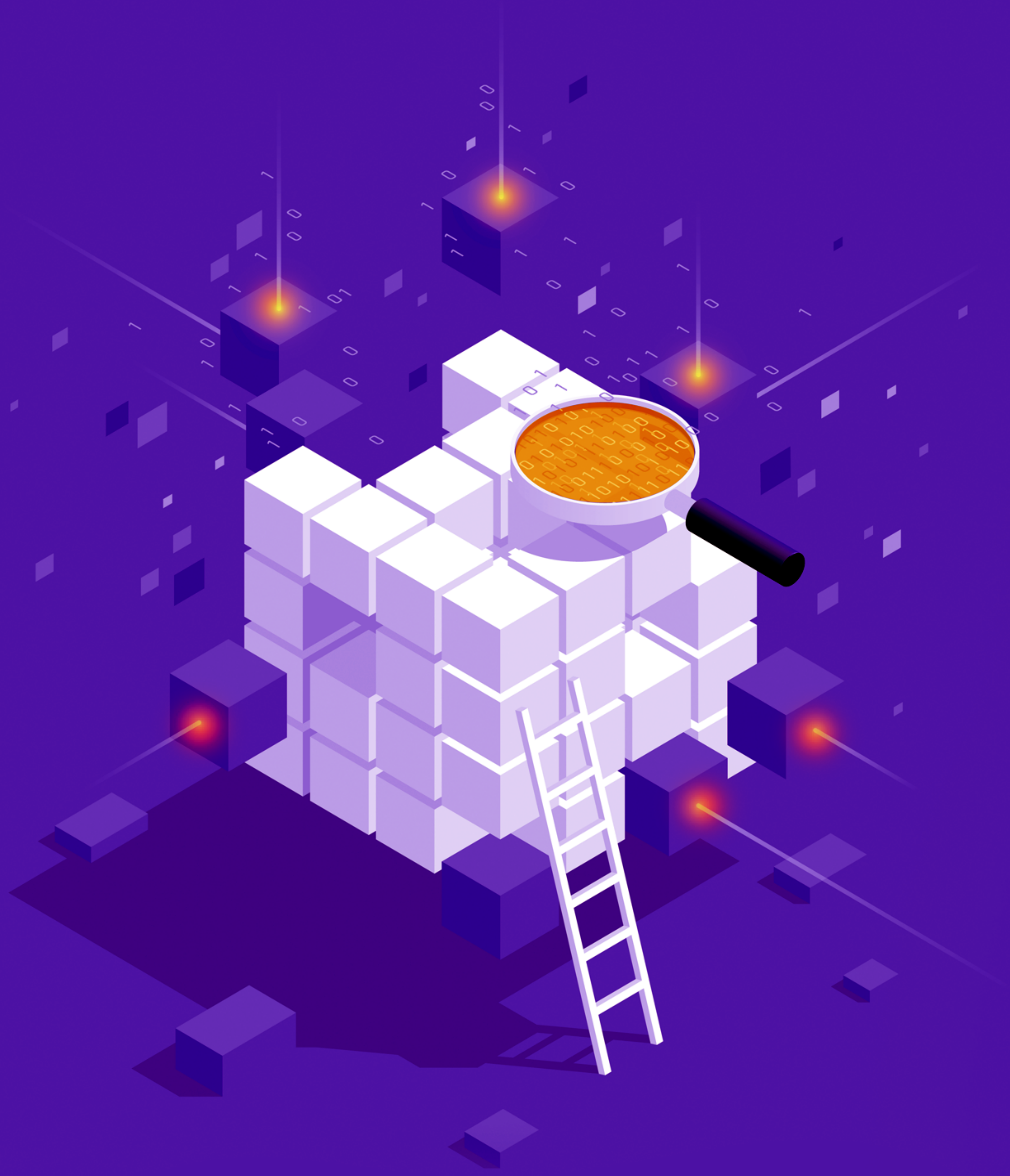
The challenges of online traffic filtering and fraud protection arose in parallel with the spread of the Internet and have gone beyond the usual solutions for online business. A significant role in fraud protection at the moment is played by device reputation and the anomalies related to it.

In JuicyScore we believe that in order to face the contemporary challenges online businesses need to develop technologies constantly and also to change the approaches, which existed for the last 15-20 years, since many of them are gradually loosing their efficiency. For the last 6 years (since 2016) we have been working on probabilistic Device ID authentication — it is a 3-rd generation Device ID beyond current privacy regulation rules. In present material we would like to shed light on some of the most effective ways of solving these problems and also to present our vision of Device ID usage trends.

The main purpose of calculating or determining Device ID remains device authentication in order to solve various applied tasks. For example, accurately determined Device ID helps to prevent repeated applications from a device with signs of fraud risk. The Device ID is quite an effective tool, at least for the reason that changing the real Device ID is tantamount to buying a new device, which leads to additional financial and operational barriers for fraudsters or toxic users.

It would seem that accurate determination of a Device ID is a fundamental tool in the fight against the risk of fraud and should be a fairly simple solution, but there are a number of serious technical aspects that need to be considered:

- **Device ID verification accuracy.** This problem is related to the fact that practically every device can be modified in terms of factory settings and authentication details or the parameters, relevant for Device ID authentication;
- **Online privacy.** This aspect is connected with the fact that for further sustainable development of the biggest tech companies in terms of digitisation users privacy solution is highly required. At the moment we can see that the problem of “passive privacy” has been solved to some extend - in order to do that different measures have been taken, to name a few, MAC-addresses routers removal from public network, third-party cookies, Google Privacy Initiative (https://privacysandbox.com/), Apple Privacy Requirements, FireFox third-party cookies initiative and some others. It is quite obvious, that digital risk management technologies should be in line with trend of privacy requirements development.



JuicyScore

1st Generation Web ID/ Device ID

ID based on virtual user data

Device fingerprint of the first generation may not be regarded as a device fingerprint in the modern sense. Fundamentally it is based on the characteristics of the device, but to a greater extent on the traces user leaves on the Internet - to name a few the most common, email or mobile phone number. This also may be a reversible hash and in some cases even a clear text.

The main flaws of such fingerprinting were:

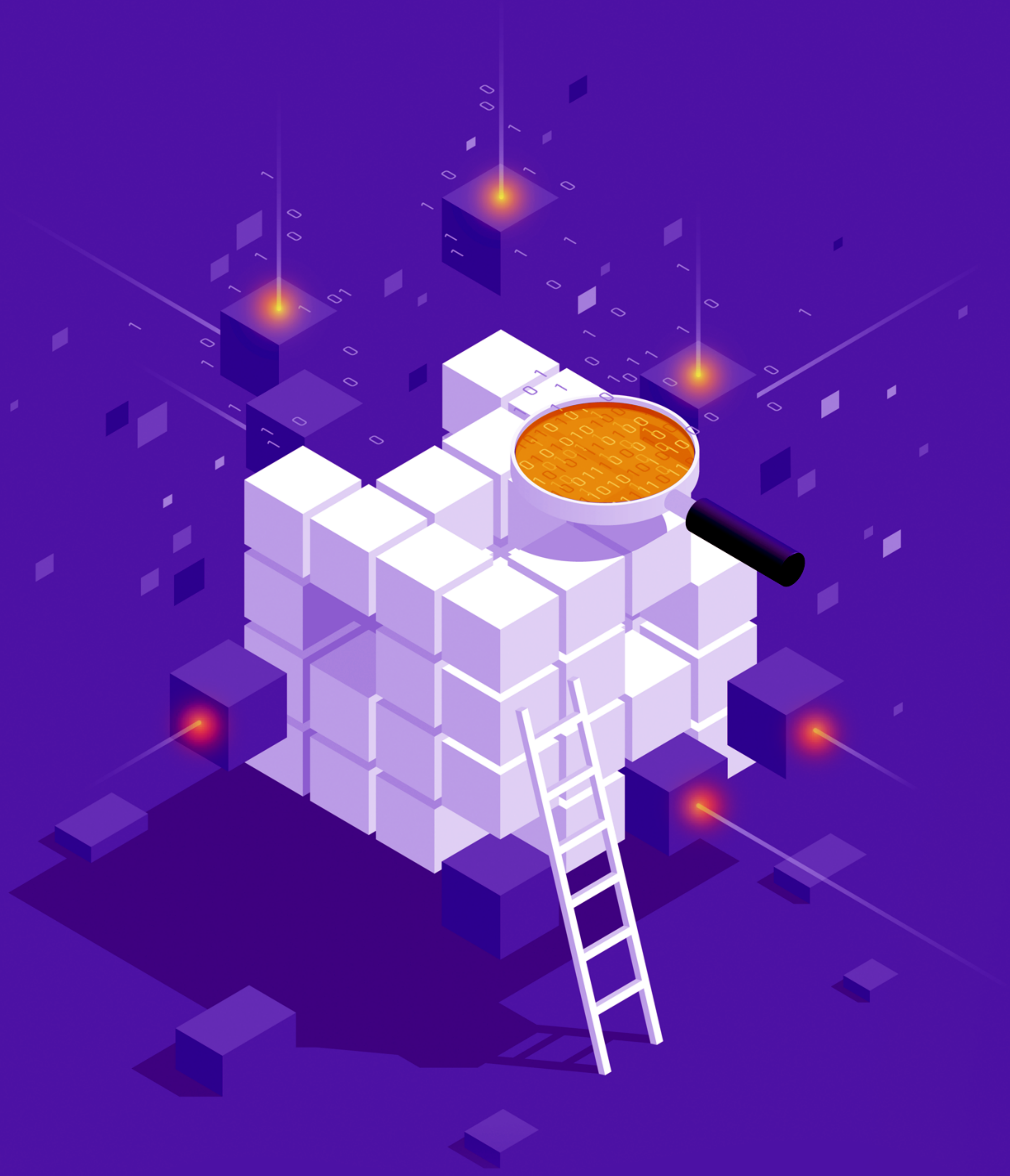
- Real user data compromise risk;
- The ease of manipulation - a user can create and operate with an infinite number of different email addresses and a significant number of different phone numbers in various combinations. In addition, forging several fields of personal data is not such a difficult task for an unscrupulous online business client;
- This generation of Device ID was created without taking into account modern privacy requirements.

All these reasons created a precondition for other technologies development on the Internet.

JuicyScore

2nd Generation Web ID/ Device ID

Device statistical ID



Within this generation of IDs, which began to develop actively in the last 10-15 years, we distinguish IDs related to statical device data (MAC address, EMEI, browser hash (browser hash is often called Device ID), persistent sessions and etc.). Let's take a closer look at some of them:

Browser Hash

This fingerprint detection technology is based on the analysis of statistical components of browsers such as model, browser and operating system versions, system language, screen resolution, time zone, clock data down to millisecond as well as a list of standard fonts installed on the device. Most of the methods are available for all browser modes, as they are necessary for the browsers normal operation. Among the various directions of this method we distinguish classical device fingerprinting, canvas fingerprinting, webgl fingerprinting, audio fingerprinting and a number of others.

The main problem is instability, as browser hash masks change when manipulating a small number of the parameters above as well as in case of browsers changing.

The second important problem is low accuracy, since the browser hash itself does not allow you to authenticate the device with a high degree of probability.

Persistent sessions

One of the first device detection technologies was Evercookie or Persistent cookie technology. Its essence lies in the fact that this type of cookie does not just store information in one data warehouse, such as an http cookie, but uses all available storages of modern browsers - modern HTML 5 standard, Session Storage, Local Storage and others. The ETag header is also used - this is an http header, very short, but you can encode any information in it, and if Java is installed, Java presistence API is used. In addition to this, the PNG Cookies mechanism is used (encoding a piece of information as a small PNG file) and playbacks it through the Canvas API.

Despite all this, persistent-cookies methods practically do not work under an incognito or in private modes of all the contemporary browsers.

Classic Device Fingerprinting

Code libraries checks user's browser for all specific and unique settings and data for this browser and the device as well, the data is reassembled into a string and cached by a certain algorithm. The basis of hash key formation is UserAgent. Browser language, time zone (offset from UTC), fonts, color palette, separate platform-dependent constants and other data specific to the user and platform are added to it, sessions of Internet giants can also be used. Information about plugins installed as well as all multimedia types or main types that support this plugin is used in order to increase uniqueness. Canvas Fingerprinting technologies are also used - certain text is drawn on a hidden canvas element with certain effects applied to it. And then the resulting image is serialized into a byte array and converted to base64.

The main flaws of this approach include the frequent changeability of the UserAgent in modern browsers, specific features of hardware implementation of devices by some manufacturers as well as the features of outdated browsers, lack of integration with Flash and Silverlight and some others.

According to a study by the Electronic Frontier Foundation (a part of the Panopticlick project, https://coveryourtracks.eff.org), the uniqueness of a fingerprint is about 90-94%. Due to the collection and analysis of a large number of parameters and settings of device hardware and software, it's possible to ensure the necessary level of entropy and uniqueness of each digital fingerprint. Existing fingerprinting-based device detection technologies have caused a significant breakthrough in terms of incoming traffic assessment, web resource audience assessment etc.

The main problems of classical fingerprinting are related to the following aspects:

Fingerprinting relies on a limited set of root parameters, the number of which is decreasing over the time (browsers', marketplaces and software platform developers policies tightening as well as new regulatory requirements). For example, reducing the readable UserAgent field has resulted in a significant reduction in the effectiveness of a number of fingerprinting technologies. Reducing the set of initial available parameters for device fingerprinting inevitably leads to an increase in collisions.

Digital fingerprint calculation technology. Basically, there are two approaches: front-end and back-end fingerprinting. On the front end it can be easily assembled and read. The downside is that the scammer will also be able to read and make changes to the Device ID, which seriously reduces the potential toolkit for building other ways to identify and match devices, while the Device ID calculation technology on the back end is not visible at all.

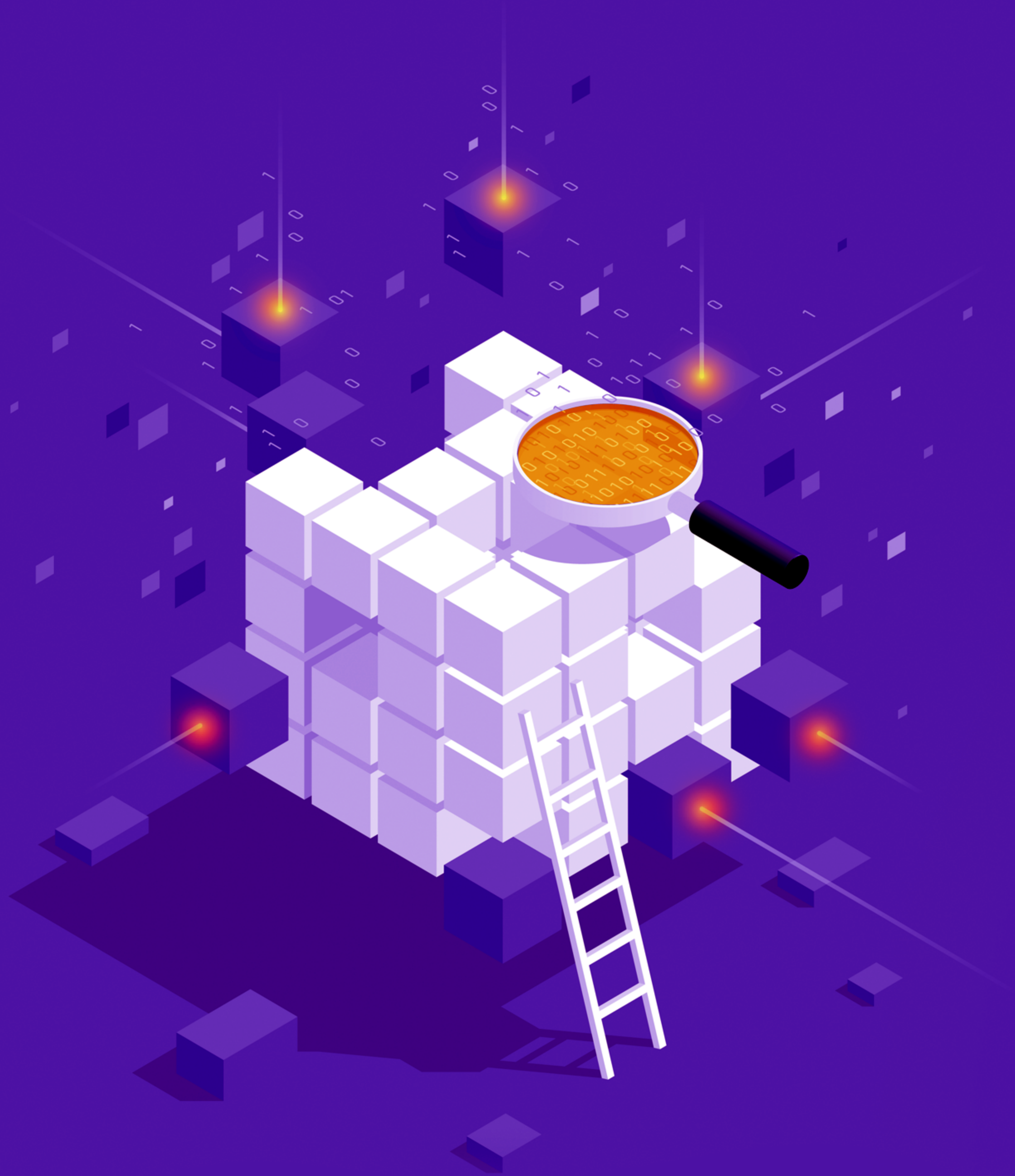
Device fingerprint stability regarding the same real device (especially if we are talking about a trustworthy client) is not sufficient for the effective use of these technologies in industries with higher accuracy requirements.

In an extension of the above problem, the opportunity to use such fingerprint in account-centric systems as well as to calculate any attributes in relation to this accounting unit and use such data for decision making may sometimes be limited.

JuicyScore

3rd Generation Web ID/ Device ID

Probabilistic Device ID. Nothing personal



Our solution is based on the following approaches that allow us to build a more robust probabilistic device ID: JuicyDeviceID with a uniqueness level of 95-99%+, depending on the required probability setting, service response time tolerance / JuicyDeviceID calculation time and various secondary aspects of international online markets (density and quality of the Internet infrastructure, most common devices of a certain model etc.):

- A wide range of additional methods;
- Server-network methods: Juicy_TCP/IP fingerprinting, AI_TLS and a number of other technologies;
- Behavioural patterns and load tests: these approaches are related to the fact that the behavioral use of devices and their performance is unique. Even the devices, which were released on the same day and on the same assembly line, will have different behavioral characteristics;
- Unique ways of building robust device fingerprints: we constantly analyze various anomalies that lead to loss of stability, discuss their “normality” with industry experts and take these aspects into account during our work;
- Artificial intelligence: we use our own proprietary artificial intelligence best practices and know-hows in order to determine to which extend two devices are similar to one another, basing on 50k+ parameters and their combinations. These methods make it possible to take into account possible anomalies on the device and provide convergence within the framework of the problem of effective device authentication;
- From this perspective we have a very fast feedback loop when new fraud patterns appear. We can look at every device that showed signs of fraud risk a month ago, find new characteristics and anomalies and analyze the exact risk of fraud as well as all the details;
- Architecture uniqueness: Our JuicyDeviceID calculation architecture will retain its uniqueness even with the mass adoption of Web3. Its main advantage is flexibility, which, in conjunction with the speed of development, shows a remarkable result;
- Significant increase in user’s privacy and security by means of removing the alternative to synchronize probabilistic device_ids at the architecture level, avoiding the use of direct user identifiers due to the limited time of use / limited time of stability of the probabilistic device_id as well as by means of mandatory informing of the users about the collection and evaluation of probabilistic device_ids .

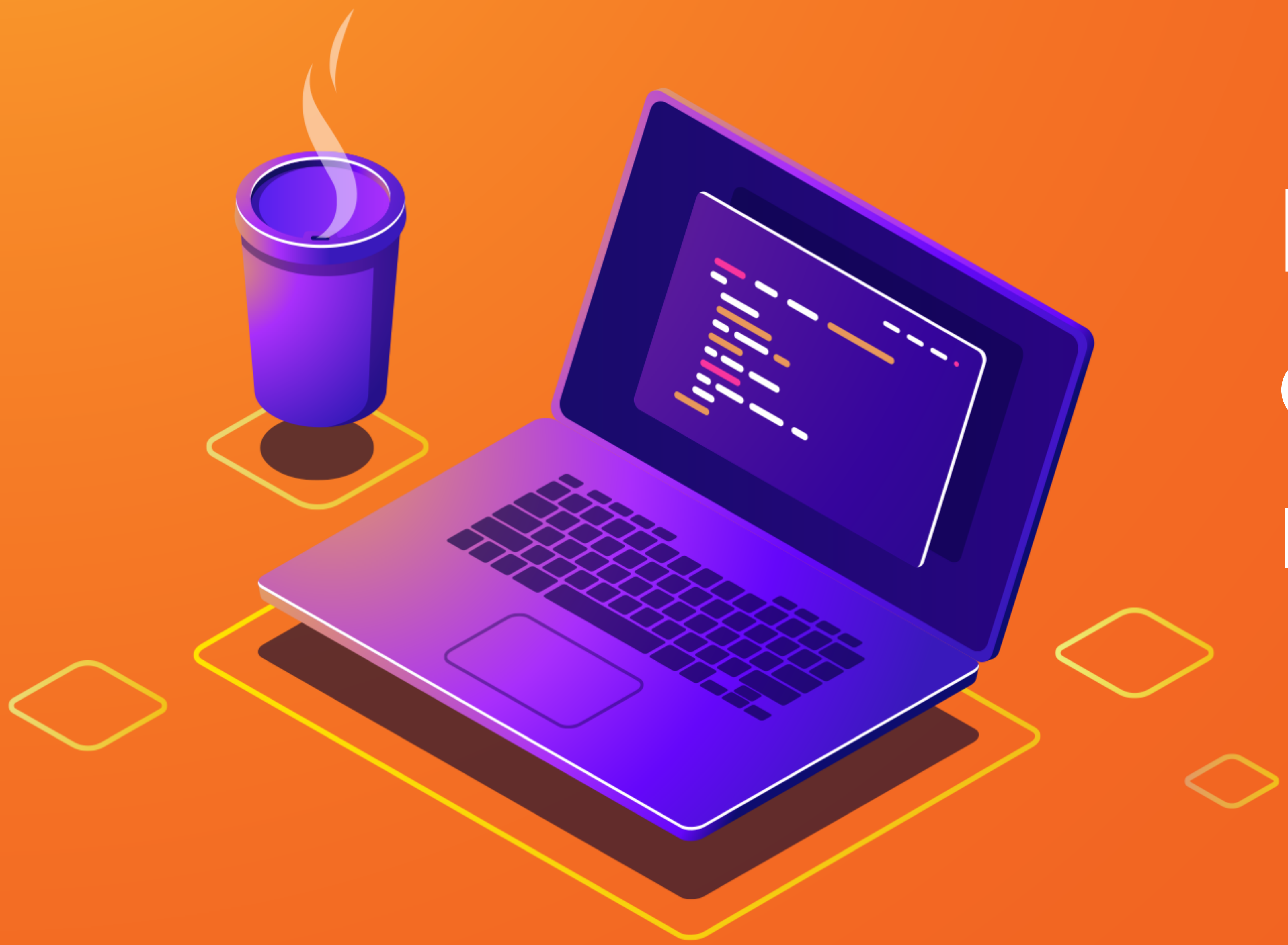
JuicyDeviceID is an effective way to solve application fraud prevention and detect multi-accounting:

- Duplicate accounts and multi-accounting risk reduction: when a fraudster uses the same device or multiple devices to create multiple personal accounts. Device ID helps to reduce risky accounts and improve unit-economy significantly.

- Protection of personal accounts of users from unauthorized access by detecting unfamiliar devices or devices with a new Device ID.

- A strong Device ID significantly strengthens the account-centric systems on the online business’ side and allows to embed such ID to the decision-making system (filters, rules, models, reporting). Different variable markers and attributes in relation to such ID have great predictive power, allow to determine the characteristics of the device, Internet connection and user behaviour and therefore to increase the quality and separating power of various models on the online business side.

JuicyScore



Device ID vs methods of virtualization and randomisation

A strong and stable ID device is a necessary, but not always sufficient tool for an effective risk management in online business. According to our experience and best practices in various markets, the most dangerous cases are associated with the use of professional device randomization methods or the so-called randomizers or anti-detect tools. Any business needs to have a set of technologies to define randomization, virtualization and remote access.

We have developed 300+ randomization detection technologies and are constantly improving our technological stack.

Any online business that works with any assets is in a great need of a stable Device ID. The most common use case for Device ID is to prevent or identify the risk of user fraud. In order to improve business performance many companies prefer to use not one, but a multiple solutions aiming to use their synergy to add more value. Protection across personal data verification, for example, such as data from the credit bureaus, telecom operators, social networks, as well as verification of digital fingerprints of devices and anomaly detection, gives the best result in terms of the synergy of these two concepts, which positively affects the final level of risk in the portfolio and ROI. Thus, companies have to admit that the use of multiple solutions plays a crucial role in growing a successful and sustainable business.

JuicyScore

The Future

Some closing thoughts

The advantages of our method are quite obvious. First of all they include security: all device calculations are performed at the server level, not at the level of the user's device. Thus, we reduce the impact of possible manipulations, preserve privacy and reduce the load on user's device. Moreover, it allows us to constantly detect various randomizers / anti-detect plugins, virtual machines as well as maintain privacy, since in this mode we can require online sites to follow the data processing policy, inform users and ensure the operation of the service only in order to prevent fraud and reduce operational risks.

Currently the most informative and successful antifraud solutions have to meet the requirements commonly accepted by the industry:

- Real time fraud risk detection: hundreds of fraudsters can attack a financial institution's resource in a short period of time;
- High data information value - in order to improve the quality of decision making systems;
- Analysis of user behaviour and verification of hidden correlations.

However, as we all know, in antifraud and risk management there is no single and universal approach that would solve any problem and give 100 percent result. JuicyScore team believes that every new self-sustainable approach will find it's place.

JuicyScore

Contact us

Protect your business and let it grow without risk

Contact us now: https://juicyscore.ai/en/ready-to-connect/