

# Криптографические подписи в данных от камеры как средство борьбы с фальсификациями

Арсений Ворошилов по материалам [techxplore.com](https://techxplore.com)

**А**I-генераторы видео появились относительно недавно, но уже получили широкое распространение. Ими пользуются как любители, которым нравится экспериментировать с новыми технологиями, так и профессионалы, которые видят в таких генераторах еще один инструмент для воплощения своих творческих замыслов. Да и жажда познания им присуща не меньше, чем любителям.

Но, как и любая технология, генерирование видео с помощью искусственного интеллекта может использоваться как во благо, так и с не очень хорошим умыслом. Порой даже с намерением фальсифицировать те или иные факты, а то и выдать за реальность то, чего на самом деле не было. Профессионалы, как правило, быстро, чуть ли не с первого взгляда, распознают подделку, а вот массовая аудитория, не всегда искушенная в визуальном определении такого рода фальсификаций, легко может стать жертвой дезинформации. К чему это может привести, долго рассказывать не надо. Вариантов множество – от дискредитации тех или иных персоналий и/или компаний до риска массовых волнений. Поэтому вопрос о том, как распознавать медиафальшивки, далеко не праздный.



*Вероятно, одно из наиболее известных фальшивых изображений на сегодняшний день – Папа Римский в пуховике. Изображение создано Пабло Хавьером с помощью AI-генератора Midjourney (фото Wikimedia Commons)*

Многие эксперты отмечают, что одним из наиболее вредоносных эффектов фальсификаций, сделанных с помощью AI-генераторов изображений и видео, является влияние на демократические процессы и подрыв доверия общества. Исследователи из ETH Zurich – Швейцарской высшей технической школы Цюриха (одного из ведущих технических университетов мира), разработали технологию изготовления чипов, позволяющую верифицировать аутентичность данных, сформированных сенсорами камер, в том числе видеоизображений и ви-

деоматериалов. Отчет о результатах работы ученых был опубликован в журнале Nature Electronics.

Предпосылкой к разработке стало то, что искусственный интеллект сейчас предельно упрощает манипулирование фотоснимками, видеоматериалами и аудиозаписями. Будь то сфабрикованные заявления, приписываемые политикам, или вводящие в заблуждение изображения из горячих точек, неоспоримым является факт, что социальные сети и онлайн-платформы уже переполнены подобного рода фальшивками. Последствия для общества и демократии очень серьезны – растет число людей, которые становятся жертвами таких подделок, и многие из них перестают доверять даже проверенным, заслуживающим доверия источникам информации.

Исследователи из ETH Zurich разработали технологию производства сенсоров, позволяющую решить эту проблему. Концепция предусматривает внедрение в изображения, видеоматериалы и аудиосигналы криптографической подписи. Делается это непосредственно в чипе сенсора в определенный момент фиксации контента. Эта подпись позволяет удостовериться, что данные действительно поступают с камеры или устройства записи, показывает, когда была сделана запись, и гарантирует, что она не является подделкой.

«Если данные снабжены подписью в момент их создания, любые дальнейшие манипуляции оставляют следы, – объясняет Фернандо Кардес, участвовавший в разработке технологии. – Чтобы осуществлять какие-то операции с данными, необходимо физическое вмешательство в чип, а это требует огромных технологических усилий, так что массовое создание поддельного контента для социальных сетей было бы практически невозможным». Кардес является научным сотрудником, работающим под руководством профессора инженерии биосистем Андреаса Йерлемана на факультете науки и техники о биосистемах (BSSE) в Базеле.

Технология действительно выглядит многообещающей. Особенно в свете того, что AI-генераторы изображения неуклонно совершенствуются. Несложно заметить, что движения живых объектов в кадре становятся более естественными, все реже возникают ошибки типа шести пальцев на руке, лучше поддерживается целостность видеоряда от кадра к кадру. Важно, что разработчики технологии предусмотрели и простой способ верификации с помощью публичного регистра. Это не допускающая изменений база данных – своего рода блокчейн, куда помещаются сгенерированные сенсором криптографические подписи. За сохранение подписей в публичном регистре отвечают производители камер. Такой подход позволил



*Принцип работы технологии: происходящее в реальном мире событие (1) снимается камерой, чип сенсора которой формирует и данные изображения, и криптографическую подпись в момент съемки (2). После сохранения в публичном регистре (3) подпись может в дальнейшем использоваться для подтверждения того, что запись аутентична и не подвергалась изменению (4). (Создано с помощью AI Феликсом Франке из ETH Zurich).*

бы любому верифицировать аутентичность данных в случае возникновения сомнения. Верификация должна быть доступна в любое время и выполняется путем сравнения внедренной чипом подписи, которая хранится в базе данных, с исходными данными и подтверждения источника этих данных.

*«Таким образом, вряд ли имеет значение, заслуживает ли доверия человек или технология, вовлеченные в обработку и передачу данных, – объясняет Феликс Франке, участвовавший в разработке чипа в ETH Zurich, сейчас являющийся профессором Университета Базеля. – Доверие к цифровому контенту ослабевает. Мы хотели создать технологию, которая дает людям способ проверить, действительно ли тот или иной контент подлинный».*

В принципе, эту технологию можно применить в сенсоре любого типа, равно как и в камере. В будущем социальные сети могли бы автоматически проверять подлинность контента сразу после его загрузки на платформу. Там, где это не делается, журналисты, исследователи или общественные структуры могли бы аутентифицировать контент самостоятельно с помощью простых средств.

Разумеется, чем раньше выявляется подделка или фальсификация, тем лучше. Ведь сегодня как никогда злободневно высказывание Уинстона Черчилля: «Ложь успевает обойти полмира, пока правда надевает штаны». И предложенная швейцарскими исследователями технология направлена именно на это – максимально раннюю идентификацию фальшивок. А ведь идея придать чипам сенсоров функцию внедрения криптографической подписи изначально появилась как побочный проект в лаборатории биоинженерии университета ETH Zurich. Задолго до того, как AI-системы типа ChatGPT стали объектом общественных дискуссий, лаборатория работала

над высокочувствительными сенсорами для измерения электрических сигналов от живых клеток. Междисциплинарный коллектив также обладал достаточной компетентностью для внедрения дополнительных криптографических функций непосредственно в чипы сенсоров.

*«Опасность, которую несут фальсификации, была предсказуемой», – вспоминает Франке. Поэтому еще в 2017 году рассматривался план разработки сенсора, данными с которого нельзя было бы манипулировать незаметно.*

В настоящее время созданный исследователями чип представляет собой действующий прототип и демонстрирует техническую осуществимость проекта. Требуется дальнейшие шаги, прежде чем можно будет начать коммерческое применение. Тем не менее исследователи уверены, что с помощью имеющихся на сегодня технологий и процессов чип может быть доведен до состояния работающего изделия, готового к выводу на рынок. Поэтому они уже подали патентную заявку.

*«Мы пока изучаем, как снизить цены для производителей камер и сенсоров, если они захотят внедрить новую технологию в свои чипы», – завершает Кардес.*

Что же, нет оснований сомневаться в довольно скором появлении подобных средств верификации контента. Тем более что уже есть и производители камер, создавшие что-то похожее. И это хорошо. Поскольку фальшивый медиаконтент сродни фальшивым деньгам, только наносит еще больший вред. С фальшивыми деньгами человечество бороться научилось. Да, фальшивые банкноты все еще встречаются, но довольно редко. Хочется надеяться, что технологии типа той, что разработали швейцарские ученые, поставят заслон массовому фальшивому медиаконтенту.