УДК 004.7:519.872 EDN VXVKSY

# Границы применимости формул Эрланга при расчете нагрузки контакт-центра

**Е.А. Сысоев***

ООО «Сибирская генерирующая компания», Красноярск, Россия

*E-mail: easysoev@mail.ru

**Аннотация**. В статье уточняются границы применения формул Эрланга при расчете нагрузки современного контакт-центра. Цель исследования состоит в том, чтобы показать, почему единая оценка по среднему AHT не описывает систему, где один вызов последовательно занимает SIP-линию, IVR, ASR, TTS, очередь и оператора. Материалом исследования выступает расчетная модель ПАРК, дополненная сопоставлением с классическими подходами Erlang B и Erlang C. Показано, что Эрланг остается полезным для локальной проверки однородной очереди или группы каналов, но не должен использоваться как единственный инструмент проектной оценки. Сделан вывод о необходимости многослойного расчета, в котором операторская, линейная и речевая емкость оцениваются отдельно.

**Ключевые слова:** Эрланг, контакт-центр, нагрузка, очередь, SIP-каналы, операторский слой, речевые сервисы.

**Для цитирования:** Сысоев, Е. А. Границы применимости формул Эрланга при расчете нагрузки контакт-центра / Е. А. Сысоев // Модернизация, Инновации, Прогресс (MIP ENGINEERING-VII-2026): материалы VII Всероссийской (национальной) научной конференции (Красноярск, 11–12 сентября 2026 г.). – Красноярск: Красноярский краевой Дом науки и техники Российского Союза научных и инженерных общественных объединений, 2026. – Т. 20. – С. 33-38. – EDN: VXVKSY.

# Applicability limits of Erlang formulas in contact center workload calculation

**E.A. Sysoev***

The Siberian Generating Company, Krasnoyarsk, Russia

*E-mail: easysoev@mail.ru

**Abstract.** The article clarifies the applicability limits of Erlang formulas in workload calculation for a modern contact center. The purpose is to show why a single estimate based on average AHT does not describe a system in which one call successively occupies a SIP line, IVR, ASR, TTS, a queue, and an operator. The study uses the PARK calculation model and compares it with the classical Erlang B and Erlang C approaches. The results show that Erlang remains useful for local verification of a homogeneous queue or trunk group, but it should not be used as the only design-capacity tool. The conclusion is that contact-center capacity should be calculated by layers, with operator, line, and speech-service resources evaluated separately.

**Keywords:** Erlang, contact center, workload, queue, SIP channels, operator layer, speech services.

**For citation:** Sysoev, E. A. Applicability limits of Erlang formulas in contact center workload calculation / E. A. Sysoev // Modernization, Innovations, Progress (MIP ENGINEERING-VII-2026): Proceedings of the VII All-Russian (National) Scientific Conference (Krasnoyarsk, September 11–12, 2026). – Krasnoyarsk: Krasnoyarsk Regional Science and Technology City Hall of the Russian Union of Scientific and Engineering Public Associations, 2026. – Vol. 20. – P. 33-38. – EDN: VXVKSY.

## 1. Введение

Расчет нагрузки контакт-центра обычно начинается с определения входящего потока и средней длительности обслуживания. В классической постановке этого достаточно, чтобы оценить вероятность блокировки или ожидания. Однако современный голосовой сервис уже не сводится к одной очереди. Перед оператором находятся приветствие, IVR, идентификация, распознавание речи, синтез фраз, возможное самообслуживание, очередь и callback. После разговора оператор может выполнять постобработку, а линия связи при этом уже свободна. Поэтому один и тот же клиентский вызов создает несколько разных нагрузок, которые нельзя корректно заменить одним средним АНТ.

Научная проблема состоит в том, что Erlang B и Erlang C часто применяются шире, чем позволяют их предпосылки. Эти модели описывают однородные ресурсы с понятной границей занятости, но контакт-центр с голосовым роботом имеет каскад ресурсов. Ошибка в выборе входной нагрузки приводит к двум противоположным последствиям: операторская численность может быть завышена, если все входящие попытки считать звонками к человеку, а инфраструктурная емкость может быть занижена, если не учитывать линии, ASR, TTS и повторные попытки.

Цель статьи – определить, где классические формулы Эрланга сохраняют практическую ценность, а где требуется многослойный расчет. Для достижения цели решаются три задачи: описать исходные ограничения Erlang B и Erlang C, показать различие между операторским и инфраструктурным временем занятия ресурсов, а также сопоставить классический подход с потоково-адаптивной моделью ПАРК.

## 2 Материалы и методы

Методологической основой исследования служит сравнительный анализ классических моделей телетрафика и многослойного ресурсного расчета. Историческая формула Эрланга используется как исходная модель предложенной нагрузки, где поток заявок связывается со средней длительностью обслуживания [1]. Для проверки границ применимости используются обзоры операционного управления контакт-центрами, в которых показаны роль очередей, маршрутизации, отказов, прогнозирования и WFM-процессов [2]. Дополнительно учитывается прикладная логика оптимизации контакт-центров, где Erlang C рассматривается как полезный, но частный инструмент планирования операторов [3].

В качестве расчетной схемы используется модель ПАРК. Она начинается с эффективного потока $\lambda e$, который учитывает прогноз входящих попыток, долю повторов и

поправку на пик. Затем поток раскладывается по маршрутам: к оператору, в голосовой робот и в короткое завершение. Для каждого маршрута фиксируется длительность занятия отдельного ресурса. Такой подход согласуется с современной практикой WFM, где решения принимаются на разных горизонтах планирования и должны учитывать изменчивость поступлений, расписания и маршрутизации [4].

Для сопоставления выбраны три расчетных слоя: операторский, линейный и речевой. Операторский слой рассчитывается только по обращениям, дошедшим до человека. Линейный слой учитывает время от установления соединения до разрыва вызова. Речевой слой описывает активные секунды распознавания и динамического синтеза речи. При интерпретации результатов учитывается, что мультинавыковые группы и неоднородные очереди требуют отдельного анализа качества сервиса, а не простого суммирования всех операторов в один пул [5].

## 3 Результаты и обсуждение

Первый результат состоит в уточнении предметной области Erlang B. Эта формула применима для оценки блокировки однородной группы каналов без очереди. В контакт-центре она может использоваться как контрольная нижняя граница для SIP-транка, но не как проектная емкость всей голосовой инфраструктуры. Линия занята во время приветствия, IVR, ожидания, разговора, удержания, бота и короткого завершения. Если в расчете есть исходящие кампании, callback, резерв на отказ узла и повторные попытки, проектное число каналов должно быть выше минимальной вероятностной оценки.

Второй результат связан с Erlang C. Формула полезна для проверки ожидания в операторской очереди, но только после правильного определения потока к человеку. Если все входящие попытки подать в Erlang C как операторские звонки, расчет завысит потребность в персонале. Если исключить из модели вариативность поступлений, длительность обслуживания, shrinkage и навыковую структуру, расчет занизит плановое расписание. Поэтому Erlang C должен быть встроен в WFM-логику, где теоретическая нагрузка переводится в смены, перерывы, обучение, отпуска и фактическую доступность операторов.

Третий результат показывает, что речевые сервисы не описываются операторским AHT. ASR занят активными секундами слушания, а TTS – динамическим синтезом или потоковой выдачей фразы. Если фраза заранее записана или отдана из кэша, она не должна нагружать TTS как новая генерация. Поэтому расчет речевого слоя должен учитывать не только число звонков, но и длительность бот-диалога, долю кэша, количество символов на вызов, целевую загрузку узла и резерв вариативности.

Для обобщения различий в таблице 1 приведено сопоставление трех расчетных слоев. Таблица показывает, зачем требуется разделять ресурсы: каждый слой имеет собственную единицу нагрузки, собственный источник данных и собственный управленческий риск. Такое разделение предотвращает ситуацию, когда снижение числа операторов после автоматизации ошибочно принимается за снижение всей нагрузки контакт-центра.

**Таблица 1.** Сопоставление ресурсных слоев контакт-центра.

| Слой | Единица нагрузки | Основной риск неверного расчета |
|---|---|---|
| Операторы | Обращения, дошедшие до человека; разговор, удержание и ACW | Завышение штата при включении всех входящих попыток или занижение расписания без shrinkage |
| SIP-линии | Время соединения от ответа до разрыва по всем маршрутам | Недостаток каналов при учете только разговора с оператором |
| ASR | Активные секунды слушания и распознавания | Перегрузка речевых потоков при росте доли голосового робота |
| TTS | Динамические секунды синтеза и скорость генерации символов | Недооценка узлов при игнорировании кэша и длины фраз |

Данные таблицы 1 показывают, что многослойная модель не отменяет классические формулы, а ограничивает их корректной областью. Для операторов важны поток до человека, целевая занятость и shrinkage. Для линий важна длительность удержания соединения по всем маршрутам. Для ASR и TTS важны активные секунды и вычислительная емкость. В многоэтапных контакт-центрах с переполнением и нетерпеливыми клиентами подобное разделение также требуется для корректного определения числа сотрудников и показателей ожидания [6].

Повторные попытки усиливают это различие. С точки зрения бизнеса повторный звонок может быть продолжением того же клиентского вопроса, но для телефонии он является новой попыткой, занявшей канал и сервисы маршрутизации. Исследования retrial-очередей показывают, что возвраты клиентов после занятости или отказа должны учитываться как часть системной нагрузки, иначе прогноз и последующее планирование будут искажены [7]. В модели ПАРК это отражается через эффективный поток $\lambda_e$, где доля повторов добавляется до распределения нагрузки по слоям.

Автоматизация не устраняет нагрузку, а переносит ее между слоями. При росте доли голосового робота операторов может требоваться меньше, но ASR и TTS будут работать дольше и чаще. Современные исследования влияния искусственного интеллекта на управление call centre подчеркивают, что автоматизация меняет роли сотрудников, требования к компетенциям и показатели эффективности [8]. Для расчета емкости это означает

необходимость оценивать не только экономию операторского времени, но и стоимость вычислительной инфраструктуры, кэша, мониторинга качества речи и отказоустойчивости.

Практический вывод заключается в том, что среднее значение AHT должно быть заменено набором длительностей. Для операторов нужны Ttalk, Thold и Tacw. Для линий нужны Tivr, Wq, Ttalk, Thold, Tbot и Tdrop. Для ASR и TTS нужны активные секунды по маршрутам. Если какой-либо параметр отсутствует, в расчетной форме следует фиксировать «данных нет», а не подставлять удобную догадку. Иначе модель создает ложную точность и скрывает место, где данные еще не готовы для проектного решения.

Таким образом, рациональное применение Эрланга возможно только после декомпозиции клиентской сессии. Сначала определяется эффективный поток, затем поток распределяется по маршрутам, после этого рассчитываются ресурсы каждого слоя и добавляется резерв на вариативность. В этой логике Erlang B и Erlang C становятся контрольными проверками внутри более широкой модели, а не заменой анализа архитектуры контакт-центра.

## 4 Заключение

Формулы Эрланга сохраняют научную и практическую ценность, если применять их к тем ресурсам, для которых выполнены исходные предпосылки модели. Erlang B полезен для локальной проверки блокировки каналов, Erlang C – для анализа операторской очереди. Однако современный контакт-центр является каскадом ресурсов, где один вызов последовательно занимает линию, IVR, ASR, TTS, очередь, оператора и постобработку.

Потоково-адаптивная модель ПАРК позволяет устранить смешение этих нагрузок. Она разделяет эффективный поток по маршрутам, задает отдельные длительности для каждого слоя и добавляет резерв на вариативность, пиковость и доступность. В результате расчет становится объяснимым: каждое число связано с конкретным источником данных, а каждый резерв можно проверить по фактической эксплуатации.

Главный вывод состоит в том, что нерационально использовать одну формулу для всей клиентской сессии. Корректный расчет должен отдельно оценивать операторов, SIP-каналы, ASR, TTS и отказоустойчивость. Такой подход снижает риск скрытого дефицита инфраструктуры и помогает принимать согласованные решения между WFM, телефонией, владельцами голосового робота и бизнесом.

## Список литературы


1. Erlang, A. K. The Theory of Probabilities and Telephone Conversations / A. K. Erlang // Nyt Tidsskrift for Matematik. – 1909. – Vol. 20. – P. 33–39.
2. Gans, N. Telephone Call Centers: Tutorial, Review, and Research Prospects / N. Gans, G. Koole, A. Mandelbaum // Manufacturing & Service Operations Management. – 2003. – Vol. 5, No. 2. – P. 79-141. – https://doi.org/10.1287/msom.5.2.79.16071. – EDN: YCPAQP.
3. Koole, G. M. Call Center Optimisation / G. M. Koole. – Amsterdam: MG books, 2013. – 152 p. – ISBN 978-90-820179-0-8. –
4. Koole, G. M. A Practice-Oriented Overview of Call Center Workforce Planning / G. M. Koole, S. Li // Stochastic Systems. – 2023. – Vol. 13, No. 4. – P. 479-495. – https://doi.org/10.1287/stsy.2021.0008. – EDN: FISYGS.
5. Hassan, N. A. Optimizing multi-skill call center staffing using queuing models: A study of service level / N. A. Hassan, N. M. S. Abdallah, R. A. Attwa // Journal of Applied Research and Technology. – 2024. – Vol. 22, No. 2. – P. 230-242. – https://doi.org/10.22201/icat.24486736e.2024.22.2.2255. – EDN: NWYASF.
6. Manitz, M. A fast staffing algorithm for multistage call centers with impatient customers and time-dependent overflow / M. Manitz, M. Ph. Piehl // Central European Journal of Operations Research. – 2023. – https://doi.org/10.1007/s10100-023-00883-z. – EDN: IFDXZJ.
7. Dang, T. C. Retrial queueing analysis model in call center system with balanced mixing calls / T. C. Dang, D. L. Hoang, L. C. Hoa, D. D. T. Nguyen, T. S. Nguyen // Hue University Journal of Science: Techniques and Technology. – 2022. – Vol. 131, No. 2A. – P. 21–37. – https://doi.org/10.26459/hueunijtt.v131i2A.6617.
8. Sheik Emam, S. Artificial Intelligence and the Future of Workforce Management in the Call Centre Industry / S. Sheik Emam, R. Sookdeo // Acta Universitatis Danubius. Administratio. – 2025. – Vol. 17, No. 1. – P. 193–209. – URL: https://dj.univ-danubius.ro/index.php/AUDA/article/view/3609.